

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: October 10, 2006, 14:34:58 ; Search time 10518 Seconds
(without alignments)
6283.482 Million cell updates/sec

Title: US-10-623-813-85

Perfect score: 3758

Sequence: 1 MNAVEENQASGSEGVSEAS.....MSVPVIRHLPAFLKEYFACV 689

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10623813/runat_10102006_143513_7128/app.query.fasta_1
-DB=GenEmbl -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs905h
-USER=US10623813 @CGN_1_1_6706 @runat_10102006_143513_7128 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_hvg:*
13: gb_in:*
14: gb_cm:*
15: gb_da:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	3758	100.0	2318	6	AF480164 Mus muscu
2	3758	100.0	4094	6	BC007466 Mus muscu
3	3758	100.0	4192	6	AF068625 Mus muscu

4	3716	98.9	2597	14	AY271299	AY271299 Bos tauru
5	3716	98.9	2798	14	AY271298	AY271298 Bos tauru
6	3714	98.8	2371	5	AF480163	AF480163 Homo sapi
7	3714	98.8	3005	2	CS050243	CS050243 Sequence
8	3714	98.8	3005	5	AF067972	AF067972 Homo sapi
9	3714	98.8	3604	2	CS050242	CS050242 Sequence
10	3714	98.8	4258	5	AF311856	AF311856 Homo sapi
11	3714	98.8	4294	2	DD171352	DD171352 Marker fo
12	3714	98.8	4294	5	BC043617	BC043617 Homo sapi
13	3714	98.8	4395	2	CS050241	CS050241 Sequence
14	3697	98.4	2841	2	C0720469	C0720469 Sequence
15	3653	97.2	4476	5	AB208833	AB208833 Homo sapi
16	3586.5	95.4	3014	11	AB214886	AB214886 Gallus ga
17	3249	86.5	264380	12	AC120824	AC120824 Rattus no
18	3115.5	82.9	253474	12	AC095441	AC095441 Rattus no
19	3093	82.3	3377	11	AB196917	AB196917 Danto rer
20	3036.5	80.8	235547	12	AC112586	AC112586 Rattus no
21	2963	78.8	2933	11	AB196919	AB196919 Danto rer
22	2477.5	65.9	344787	12	AC112040	AC112040 Rattus no
23	2412	64.2	2191	5	AK025230	AK025230 Homo sapi
24	2288.5	60.9	3544	11	AB196915	AB196915 Danto rer
25	2191.5	58.3	168651	5	AC009474	AC009474 Homo sapi
26	2167.5	57.7	4267	5	AF176228	AF176228 Homo sapi
27	2164.5	57.6	4135	6	AF068627	AF068627 Mus muscu
28	2164.5	57.6	4163	6	AF151974	AF151974 Mus muscu
29	2164.5	57.6	4278	6	AF151970	AF151970 Mus muscu
30	2159.5	57.5	4043	2	C0726303	C0726303 Sequence
31	2159.5	57.5	4145	5	AF156488	AF156488 Homo sapi
32	2159.5	57.5	4195	6	AF068626	AF068626 Mus muscu
33	2159.5	57.5	4223	6	AF151973	AF151973 Mus muscu
34	2159.5	57.5	4335	2	CO981515	CO981515 Sequence
35	2159.5	57.5	4335	5	AF331857	AF331857 Homo sapi
36	2159.5	57.5	4338	6	AF151969	AF151969 Mus muscu
37	2149	57.2	3630	6	BC105922	BC105922 Mus muscu
38	2148	57.2	4369	5	DD321787	DD321787 Homo sapi
39	2056.5	53.4	2491	14	AY244710	AY244710 Bos tauru
40	2007.5	54.7	6248	11	AB196918	AB196918 Danto rer
41	1943	51.7	2421	14	AY244712	AY244712 Bos tauru
42	1892	50.3	3946	6	AF068628	AF068628 Mus muscu
43	1892	50.3	3974	6	AF151976	AF151976 Mus muscu
44	1892	50.3	4089	6	AF151971	AF151971 Mus muscu
45	1887	50.2	4006	6	AY078427	AY078427 Mus muscu

ALIGNMENTS

RESULT 1
AF480164
LOCUS
DEFINITION Mus musculus DNA cytosine methyltransferase 3a2 (Dnmt3a2) mRNA,
complete cds.
ACCESSION AF480164.1 GI:23954439
VERSION
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Chen,T., Ueda,Y., Xie,S. and Li,E.
TITLE A Novel Dnmt3a Isoform Produced from an Alternative Promoter
Localizes to Euchromatin and Its Expression Correlates with Active
de Novo Methylation
JOURNAL J. Biol. Chem. 277 (41), 38746-38754 (2002)
PUBMED 12138111
REFERENCE
AUTHORS Chen,T. and Li,E.
TITLE Direct Submission
JOURNAL Submitted (31-JUN-2002) Cardiovascular Research Center,
Massachusetts General Hospital, 149 13th Street, Charlestown, MA
02129, USA
FEATURES
Location/Qualifiers

```

source
1. .2318
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129S4/SvJae"
/db_xref="taxon:10090"
/chromosome="12"
/map="12A2-A3"
1. .2318
/gene="Dnmt3a2"
148. .2217
/gene="Dnmt3a2"
/function="de novo DNA cytosine methylation"
/product="DNA cytosine methyltransferase 3A2"
/protein_id="AA040038.1"
/db_xref="GI:23954440"
/translation="MNAVFENQASGSGKVEASPPAVQOPTDPAAPTATTPPEPVG
DAGDKATKAUDEPEYEDGRGIGELVWGLRGRFSMTGRRAIVSMWGRRAEET
RWMMFGDGKFSVCEVKLPLSSFCSAHQATYNKQPMYRAIIEVLQVASSRAGKL
FPACHDSDSGSKAVEVONKQIEMALGGFQPSGKGLPEPEEKNPKEVYTDMMV
EEAAAYAPPAPPAKPKSTTEKPKYKEIIDERTERLYEVROKCRNIEDICISGS
LNVLEHPLFIIGMCONKCNFLCAYOYDDGYOSCTICCGREVLMGNNCCRC
FVCEVDLLVGGAAQAIKEDPMNCGMHGTYGLARRDMPSRMQFPANNHDO
EDDPKVPYPPVAERKRPTRVSLDPGATGLVKDLSIOVDRYIASFVCEDSITVG
MVRHQKIMYVGDVRSVTKHIOEMGFPLVIGSGFCDLSIYNPARKLYEGTGRLP
FEFRLHDPARKEGSDRPFVWLFEENVAMGVSDKRDISRFLSESNVMTDAEVSAAH
RFRYEWGNLPGMNRPLASTVNDKLELOECLEHGRIAKFSKVTITITRSNIOGDQH
FVPMNEKEDILMCTEMERVFGFPHVYTDVSNMSRLARQLLGRSWSVPIRHLAPL
KEYFACV"

```

ORIGIN

Alignment Scores:

Pred. No.:	Length:	2318
Score:	3758.00	689
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	100.0%	Mismatches: 0
Query Match:	100.0%	Indels: 0
DB:	6	Gaps: 0

US-10-623-813-85 (1-689) x AF480164 (1-2318)

```

QY 1 MetAsnAlValAlGluGluGlnGlnAlaSerGlyGluSerGlnLysValAlGluGlnAlaSer 20
DB 148 ATGAATGCTGCTGGAAAGAACCGAGCCTCGAGACTTCAGAAAGGTGAGGAGGCGACAC 207
QY 21 ProProAlaValAlGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 208 CCTCTGCTGCTGACAGACCCACGACCTGCTTCCGACTGTGCGACACCCCTGAG 267
QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
DB 268 CCAATGGAGGGGATCTGGGGACAAGAAATGTAACAAAGCACCGACGATGAGCTGAG 327
QY 61 TyrGluAspGlyValArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
DB 328 TATGAGAGATGGCGGGGCTTTGGCATTGAGAGCTGCTGGGGGAAACTTGGGGCTTC 387
QY 81 SerTyrTrpProGlyValArgIleValSerTyrTrpMetThrGlyValArgSerAlaAlaGlu 100
DB 388 TCTGTGTGGCCAGCCGCAATTGTCTGTGTGATGACAGGCCGAGCCGAGCAGCTGAA 447
QY 101 GlyThrArgTyrValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB 448 GGCACCTCGCTGCTCATGTGTGTCGAGATGGCAAGTTCTCACTGTGTGTGTGGAGAG 507
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 508 CTCATCCCGCTGAGCTCTCTTCAGTGCAATCCACAGGCGCACCTTACACAAAGCAGCC 567
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerAlaGlyLysLeu 160
DB 568 ATGTACCGCAAGCCATCTACGAAGTCTCTCAGGTGCGCAGCAGCGCTGCGGGAAGCTG 627

```

```

QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysValValAlGluGlnAsn 180
DB 628 TTTCCAGCTTCCATGACAGATGAATGAAGTACAGTGCAAGGCTGTGGAAGTGCAGAAC 687
QY 181 LysGlnMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 668 AAGCAGATGATGTAATGGGCCCTCGGTGGCTTCCAGCCCTGGGGTCTTACAGGCGCTGAG 747
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTyrValGluPro 220
DB 748 CCACCAAGAAAGAAAGAAATCTTACCAAGGAAGTTTACCCGACATGTGGTGGAGCCT 807
QY 221 GluAlaAlaIleTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
DB 808 GAAGCAGCTGCTTACGCCCAACCCCAAGCAGCAAGAAACCAAGAAAGCACAACAG 867
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
DB 868 AAACCTAAGTCAAGAGATCATTTATGAGGCAACAGGAGCGGCTGTGTATGAGGTG 927
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 928 CGCCAGAAAGTCAGAAACATGAGACATTTGTATCTCATGTGGAGGCTCAATGTCAAC 987
QY 281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 988 CTGGAGCAACCACTTTATTTGAGAGCATGTGCGCAAGCTGTAGAACCTGCTTTGAG 1047
QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB 1048 TGTGCTTACAGTATGACAGATGAGATGGGTACCAAGTCTTATGACATCTGCTGGGGGG 1107
QY 321 ArgGluValLeuMetCysGlyLysAsnAsnAsnCysCysArgCysPheCysValAlGluCysVal 340
DB 1108 CGTGAAGTGCATATGTGTGGAAACAACATGCTGCAAGTGTGTTGTGTGCAAGTGTG 1167
QY 341 AspLeuLeuValAlGlyProGlyValAlaAlaGlnAlaIleLysGluAspProTyrAsnCys 360
DB 1168 GATCTCTTGTGGGGCCAGAGACTGCTCAGGCGCATTTAAGAAAGACCCCTGGAACCTG 1227
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTyrProSer 380
DB 1228 TACATGTGGGGCATTAGAGGACCTATGGGCTGCGCAAGCGGGAAGACTGGCCTTCT 1287
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
DB 1288 CGACTCCAGATGTTCTTTGCCATTAACATGACCGAATTGACCCCAAGGTTTAC 1347
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 1348 CCACCTGTGCGAGCTGAGAGAGAGAGAGGCCATCCGCTGCTCTCTTGAATGGATT 1407
QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
DB 1408 GCTACAGGGCTCTGTGTGCTGTAAGAGCCTGGGCATCCAAGTGAACGCTACCTCTCC 1467
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 1468 GAGGTGTGTGAGAGATCTTCATCAACGCTGGGCAAGTGGCGCAACAGGAAAGTCAATGAC 1527
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
DB 1528 GTGGGAGAGTCCGACGCTGACAGAGAGAGGCCATCCGCTGCTCTTGAATGGATTCCACCTG 1587
QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 1588 GTGATTGGAGGAGTCCCTGCAATGACCTTCATTTGCAACCTGCGCGCAAGGAGACTT 1647
QY 501 TyrGluGlyThrGlyValArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
DB 1648 TATGAGGGTACTGGCGGCTCTTCTTGAATTCACCGCTCTCATGATGACGGGGCCC 1707
QY 521 LysGluGlyAspAspArgProPhePheTyrPheGluAsnValValAlaMetGlyVal 540

```

```

Db      |||1708 AAGGAGGAGATGATGCGCCCTTCTTGCTCTTTGAGAGATGTTGGCGATGGCGCTT 1767
Qy      541 SerAepLyArGAspIleSerArGpHeuLeuGluSerAsnProValIleTlleaPaLaLyS 560
Db      1768 AGTGACAAAGAGGACACTCTCGCATTTCTTGACTTACCCCTGTGATGATTCAGCCCAAA 1827
Qy      561 GluValIseAlaIahIsArGValArGTrIleThrThArGSeArAsnSerIleLyGIngly 580
Db      1828 GAAGTCTGCTGCACACAGGGCCCTTACTTCTGGGGTAACTTCTTGACATGAACAGG 1887
Qy      581 ProLeuAlaSerThrValAsnAspLySLeuGluLeuGInGluCySeuGluInIleGlyArG 600
Db      1888 CTTTGGCATCCACTGTGAATGATTAAGCTGGAGACTCCAAAGAGTGTGGAGACCGGCACA 1947
Qy      601 IleAlaIyPheSerLySValArGTrIleThrThArGSeArAsnSerIleLyGIngly 620
Db      1948 ATAGCCAAAGTCAAGCAAACTGAGACCATTTACCACACAGGTCAAACTCTTAAAGCAGGCG 2007
Qy      621 LySArpGInHIsPheProValPheMetAnGluValAspIleLeuTrpCySthrgIu 640
Db      2008 AAAGACCAACATTTCCCGCTTCTGAAAGAGAGAGAGACATCTCTGTGTGCTACTGAA 2067
Qy      641 MetGluArValPheGlyPheProValHIsTyThrAspValSerAsnMetSerArGLeu 660
Db      2068 ATGGAAGAGGTGTTGGCTTCCCGCTTCACTACAGACAGTCTCCCAACATGAGCGCTTG 2127
Qy      661 AlaArGInArGLeuLeuGlyArGSeArTrpSerValProValIleArGHisLeuPheAla 680
Db      2128 GCGAGGCAAGAGACTGCTGGCGGCGATGCTGAGAGCGTCCGCTCATCCGACACTTTCGCT 2187
Qy      681 ProLeuLySGLuTyThrPheAlaCySVal 689
Db      2188 CCGCTGAAGAAATTTTCTTGTGTG 2214

RESULT 2
LOCUS   BC007466                4094 bp    mRNA    linear    ROD 29-JUN-2004
DEFINITION Mus musculus DNA methyltransferase 3A, transcript variant 1, mRNA
            (cDNA clone MGC:5662 IMAGE:3492853), complete cds.
ACCESSION BC007466
VERSION   BC007466.1 GI:13938620
KEYWORDS  MGC.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
          Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4094)
          Strusberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
          Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,
          Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K.,
          Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
          Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
          Stalczer, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
          Carninci, P., Frange, C., Raha, S.S., Logucliano, N.A., Peters, G.J.,
          Abramson, R.D., Muliahy, S.J., Bosak, S.A., McEwan, P.J.,
          McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
          Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,
          Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriques, S.,
          Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
          Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
          Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
          Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,
          Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
          12477932
          Strusberg, R.

```

TITLE

JOURNAL

Direct Submission
Submitted (01-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpax11.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAX Plate: 8 Row: n Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6661208.
Location/Qualifiers

FEATURES

source

gene

CDS

```

1..4094
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech IT"
/db_xref="taxon:10090"
/clone="MGC:5662 IMAGE:3492853"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MMTV."
/clone_lib="NCI-CCAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..4094
/gene="Dnmt3a"
/db_xref="GeneID:13435"
/db_xref="MGI:1261827"
245..2971
/gene="Dnmt3a"
/codon_start=1
/product="DNA methyltransferase 3A, isoform 1"
/protein_id="AA07466.1"
/db_xref="GI:13938621"
/db_xref="GeneID:13435"
/db_xref="MGI:1261827"
/translation="MPSGPGDTSSTSSLEREDRKEGEDEENRKEERQEPATARK
VGRGRKRKHPPVSSSDTRPDPAVTTKQPMADSGPDDLPGDLERKSEEPQEGS
PAQGGKGAAPAGEGTETPPRPAASAVENCGVTEGRASAGEGEOKPTNISKME
GSRGLRGGLWESSLRQPMRPLTFQGDPIYISKRDENLANKBAEKAKVIA
VMNAVEENQASGEQKEASPPAVQPTDPASPVATTPPEVGDDADKATKAD
EPEYEDGRGFGIGELVWKLRFQSWPGRIVSWMTGSRABEGRWVWFGDGKFSV
VCVEKLMPLSFECSFHOATYKOPMYKAIYELVQVSSRAKGLFPACHSDSDSG
KAVEENKQMIEMALGFQSPGKLEPEERKNPKYVDMDVPEAPAAVAPPPA
KKPRSTKTKKVKKIEDEPRRLYVVRCKRNICISGSLANTLIEPLTIGG
MKNQKNCFLCAOYDDDGQSYTTCGGRVILMCNNCKRCFCVPCUJLYGCG
AAQQAIXKDPNNCTWCGKGTGYGLRRREDPSPRLOMFANNDDEFPDPKYPPVPA
EKKRIRVLSLFDGATLVLKODGIVDRIARVEVEDISTYGMVHQKIMVGD
VRSYVQKIQWGPEDLVIGSPCDISIVPARGLVLEGGRGLFEFRYRLHDAKPK
EGDDPFPFVLPENNVAMGVSDKRDISRELSNPVMI DAKEVSAARARVFMGLDGM
RLASTVNDKLEQCELEHGRILAKSKYRTITRENSIKQGDQHFVPMNKEDILW
CTEERVYGFVHTYDVSNMRRLARQLLGRSWSVPVIRHLPALKEFACV"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 0
Score: 3758.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 6

```

```

Length: 4094
Matches: 699
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

```

```

US-10-623-813-85 (1-689) x BC007466 (1-4094)
OY 1 MetAsnAlaValGluGlnAsnGlnAlaSerGlyGluSerGlnLysValGluGlnAlaSer 20
DB 902 ATGAATGCTGTGGAGAGAACCAAGCCCTGTGGAGATCTCAGAGAGTGGAGGAGCCAC 961
OY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 962 CCTCCGCTGTGCAGACGCCACGAGCCCTGCTTCCAGCTGTGCCACCACTCCGTAG 1021
OY 41 ProValGluGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
DB 1022 CCAATAGAGAGGGGATGCTGGAGACAAAGATGCTTACCAAGACACCCAGCATAGCTTAG 1081
OY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTropGlyLysLeuArgGlyPhe 80
DB 1082 TATGAGGATGGCCGGGGCTTTGGCATTTGGAGAGCTGTGTGGGGGAACTTCGGGGCTTC 1141
OY 81 SerTTPTPProGlyArgIleValSerTTPTPMetThrGlyArgSerArgAlaAlaGlu 100
DB 1142 TCCTGTGGCCAGCGCCGAATGTGTCTGTGTGATGACAGCGCGAGCCGAGACCTGAA 1201
OY 101 GlyThrArgTropValMetTropPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB 1202 GGCACCTCGGTGGGTCTGTGTGTGGAGATGGCAATGTCAGTGTGTGTGTGGAGAG 1261
OY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 1262 CTCATGCGCTGAGCTCTTCTGTGACGTGATTCACAGGCCCTACAAACAGCAGCCC 1321
OY 141 MetTyrArgValAlaIleTyrGluValIleGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 1322 ATGTACCCGAAAGCCATCTACGAAGTCTTCAGAGTGGCCAGCAGCCGTCCGGAAAGCTG 1381
OY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180
DB 1382 TTTCCAGCTTGCCTACAGCTATGAAAGTGAAGTGAAGTGGCAAGCTTGGAAAGTGCAGAAC 1441
OY 181 LysGlnMetIleGluTropAlaLeuGlyIleGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 1442 AAGCAGATGATTAATGAGGCCCTCGGTGGCTTCACCCCTCGGTCTTAAAGGCCCTGGA 1501
OY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTropValGluPro 220
DB 1502 CCACCAAGAAAGAGAGAAATCTTACAGGAAGTTTACACGACATGTGGGAGGACCT 1561
OY 221 GluAlaAlaAlaTyrAlaProProProAlaLysLysProArgLysSerThrThrglu 240
DB 1562 GAAGCGCTGCTTACGCCCAACCCCAACGCAAAACCCAGAAAGAGCAACAGAG 1621
OY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
DB 1622 AAACCTTAAGGTCAAGAGATCATTTGATGAGCGCACAAAGGAGCGGTGTGTGAGGTG 1681
OY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1682 CCCCAGAAATGCAAGAAATCGAGGACATTTGATATTCATGTGGAGCCCTCAATGTCCACC 1741
OY 281 LeuGlnHisProLeuPheIleGlyIleGlyMetCysGlnAsnLysLysAsnLysPheLeuGlu 300
DB 1742 CTGGAGACCCCACTTCTTATTTGAGAGCATGTGCCAAGACTTAAGAACTCTTCTTGGAG 1801
OY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB 1802 TGTGCTTACAGATGACAGATGAGTGTGAGTCAAGTCTTATTTGACCATCTGTCTGTGGGG 1861
OY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
DB 1862 CGTGAAGTCTCATGTGTGGAGCAACAACTGCTGAGGCTTGTGTGTGTGTGTGTGTGTGT 1921
OY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTropAsnCys 360
|||||
DB 1922 GATCTCTGTGGGGCCAGAGACTGTCTCAGGACGCCATTTAAGAGAACCCCTGAACTGC 1981
OY 361 TyrMetCysGlyHisLysGlyTyrTyrGlyLeuLeuAlaArgArgGluAspTropSer 380
DB 1982 TACATGTCCGGGCAATMAAGGACCCTATGGGCTGTGCAGAAAGCGGAAAGACTGGCTTCT 2041
OY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
DB 2042 CCACTCCAGATGTTCTTTTGCATTAACATGACCGAAGATTTGACCCCAAGTTTAC 2101
OY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 2102 CCACCTGTGCCAGCTGAGAAAGAGAACCCATCCCGTGTGTGTCTTGTATGGGATTT 2161
OY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAlaAspArgTyrIleAsp 440
DB 2162 GCTACAGGAGCTCTGTGTGTGTAAGGACCTGGGACATCCAAAGTGGACCTTACATTCCTC 2221
OY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 2222 GAGTGTGTGAGACTTCATACGCTGGGCAATGTGCGGACCAAGGAAAGATCATGTAC 2281
OY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
DB 2282 GTGGGAGAGTCCGACGCTCACAGAAAGCATATCCAGAGATGGGGCCCATTCGACCTCG 2341
OY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAspProAlaArgLysGluLeu 500
DB 2342 GTGATTTGAGAGGAGTCCCTGCAGATGACCTTCATGTGTACCCCTGCGCAAGGACCTT 2401
OY 501 TyrGluGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
DB 2402 TATGAGGATCTGGCGGCTCTTCTTGAAGTTCTACCGGCTCTGCATGATGCGGGCCC 2461
OY 521 LysGluGlyAspAspArgProPhePheTropLeuPheGluAsnValValAlaMetGlyVal 540
DB 2462 AAGGAGGAGATGATGCGCCCTTCTTGTGGCTTGTGAAAGTGTGTGGCCATGTGGCGCTT 2521
OY 541 SerAspLysArgAspIleSerArgPheLeuGlnLysSerAsnProValMetIleAspAlaLys 560
DB 2522 AGTGACAAAGGAGGACATCTCGGATTTCTTGAGTCTTAAACCCGTATGATTAACCCCAA 2581
OY 561 GluValSerAlaAlaHisAspAlaArgTyrPheTropGlyAsnLeuProGlyMetAsnArg 580
DB 2582 GAAGTGTCTGCTGCACACAGGCGCGTTACTTCTGGGGTAACTTCTGTGCATGAACAG 2641
OY 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnLysCysLeuGlnHisGlyArg 600
DB 2642 CTTTGGCATCCACGTGTATGATATGAGCTGTGAGAGTGTCTGAGACAGCGCAGA 2701
OY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
DB 2702 ATAGCCAAAGTTCAAGAAAGTGAGACATTTACACAGGCTCAAATCTTATAAGCAGGCG 2761
OY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
DB 2762 AAAGACACAGATTTCCCGCTCTTCAAGAAAGAGAGAGATCTGTGTGTGTGCACTGAA 2821
OY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
DB 2822 ATGGAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2881
OY 661 AlaArgGlnArgLeuLeuGlnGlyArgSerTropSerValProValIleArgHisLeuPheAla 680
DB 2882 GCGAGGCGAGAGCTGTGGGCGGATCGTGTGAGCGTGTGTGTGTGTGTGTGTGTGTGTGT 2941
OY 681 ProLeuLysGluTyrPheAlaCysVal 689
DB 2942 CCGCTGAAGAGATATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2968
|||||
RESULT 3
AF068625 AF068625 4192 bp mRNA linear ROD 06-DEC-1999
LOCUS

```

DEFINITION Mus musculus DNA cytosine-5 methyltransferase 3A (Dnmt3a) mRNA, complete cds.

ACCESSION AF068625

VERSION AF068625.2 GI:6449467

KEYWORDS

SOURCE

ORGANISM Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS 1 (bases 1 to 4192)

TITLE Okano, M., Xie, S., and Li, E.

CLONING AND CHARACTERIZATION OF A FAMILY OF NOVEL MAMMALIAN DNA (CYTOSINE-5) METHYLTRANSFERASES

JOURNAL Nat. Genet. 19 (3), 219-220 (1998)

PubMed 9662389

REFERENCE

AUTHORS 2 (bases 1 to 4192)

TITLE Xie, S., Okano, M., and Li, E.

DIRECT SUBMISSION

JOURNAL Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street, Charlestown, MA 02129, USA

REFERENCE

AUTHORS 3 (bases 1 to 4192)

TITLE Okano, M., Chijiwa, T., Sasaki, H., and Li, E.

DIRECT SUBMISSION

JOURNAL Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street, Charlestown, MA 02129, USA

REMARK

COMMENT Sequence update by submitter

FEATURES

source

1..4192

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/chromosome="12"

/map="4.0 cM"

1..4192

/gene="Dnmt3a"

1..4192

/note="Dnmt3a"

/note="similar to EST sequences deposited in GenBank"

Accession Numbers AA052791, A111043, A154890, AA240794, A756653, W58898, W59299, W91664, and W91665"

217..2943

/gene="Dnmt3a"

/function="de novo DNA methylation"

/note="contains Cys-rich region"

/codon_start=1

/product="DNA cytosine-5 methyltransferase 3A"

/protein_id="AAC4017.2"

/db_xref="GI:6449468"

/translation="MPSGPGGTTSSSLEREDDRKGEQENRCKEKEQESATARK VGRPRKRHPVESDTPKDAVTTKSPMAQDSGSPDLNGLDLEKSEEQPEGSG PAQCGKGAPEGEGTEPPEASRAVENCCVTKEGRSAGSEKOKTINIESMME GSRGLRGGLGESSLRLRPMPRLTFOADPYLISRKEDENLARKRAEKAKAVIA VNAVEENASGESOKVEASPPAVOQDPASPTAATPEPVGAGWGNKTKAAD EPEYEDGRGFGIGELVMKLRGFSMTPGIIVSWMTGRRAAGTRMFGDGKRSY EVVERKLPSTSCSAFHQATVNRKQMPYRAIVLVQVASSRAGKLPACHDSDSDG KAVEVONKOMIMWALGGFQSPGKLEPEEENPKYEYTTMWPEEAAVAAPPA KPRKSTETKPYKEIIDERTERLVYEVROKRNIEDICISGLNVLRLHPLFGG MCQNGKNCLECAVOYDDGYOYCTICGGRVLMGNCCRCFCVCVLLVPGG AAQAAIKEDPMNVCMGHKGTGLLRREDWSPRLQMFANNDQFDPKPYAPVA EKRKRTVLSTLDGATGLLVLKDLQIQVDRITASEVCDSTTGVNRRQCKIMVYG VASVQKHQIEMGPDLVIGSPCNDLSVNPARKGLVAGTGLPFEYRLHDAAPK EBDDBPFWLFFENVAVGSDKDISRFLSNPMDADEVSAHARFAPWNLPEMN RLASTVNDKLEOCLBHGRIAKFSKVTITTRNSIIGQDOHFPVPMNKEDILM CEMERVFEPHYTVDVSNMSRLARQLRGSRMSVPIRLHPLKEYFACV"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 4192

Score: 3758.00 Matches: 689

Percent Similarity: 100.04 Conservative: 0

Best Local Similarity: 100.04 Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 6 Gaps: 0

US-10-623-813-85 (1-689) x AF068625 (1-4192)

QY 1 MetAsnaIaValGluGluAsnGlnAlaSerGlyGluSerGlnLysValGluGluAlaSer 20

DB 874 ATGAATGCTGTGGAAGAGAACAGAGCTCTGTGAGAGTCCAGAGGTGAGGAGGCGAC 933

QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThValAlaThrThrProGlu 40

DB 934 CCTCTGCTGTGCACACGACCGACCTCTGCTTCCACGTGTGCCACACCTTGAG 993

QY 41 ProValGlyGlyAspAlaGlyAspLysAsnaIaThrLysAlaAlaAspAspGluProGlu 60

DB 994 CCAAGTAGAGGAGGAGTCTGGGACAAAGATGCTACCAAGCAGCCGACATGAGCTGAG 1053

QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80

DB 1054 TATGAGGATAGGCGGGGCTTGGCATTTGGAGAGTGTGTGGGGGAACTTCGGGCTTC 1113

QY 81 SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyArgSerArgAlaGlu 100

DB 1114 TCCTGTGCGCCAGCGCGCAATTGTCTTGTGTGATGACAGCGCGGAGCGAGCAGCTGAA 1173

QY 101 GlyThrArgTyrValMetThrPheGlyAspGlyLysPheSerValValCysValGluLys 120

DB 1174 GGCACTCGGTGGGTCATGTGTGTCGAGATGGCAAGTCTCTAGTGTGTGTGAGGAA 1233

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140

DB 1234 CTCAATGCGCTGAGCTCTCTTGCAGTGTCTTCCACAGGACCACTCAACAAAGCAGCC 1293

QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160

DB 1294 ATGTACCGCAAGCCATCTACCAAGTCTCCAGGTCGAGGAGCGGTCCGGGAGCTG 1353

QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180

DB 1354 TTTCAGCTTGCACATGACATGATGAAAGTGCAGTGGCAAGCTGTGAAAGTCAGAAC 1413

QY 181 LysGlnMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200

DB 1414 AAGCAGATGATGAATGGGCGCTCGGTGCTTCCAGCCCTCGGTCCTAAGGCGCTGAG 1473

QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTyrValGluPro 220

DB 1474 CCACCAAGAGAGAGAGATCTTACAGAGATTACACGACATGTGGGTGGAGCT 1533

QY 221 GlnAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240

DB 1534 GAAGGAGCTGTTAGCCGCCACCCACCCACGACGCAAGAACCCAGAAAGCACAACAG 1593

QY 241 LysProLysValLysGluIleLeuAspGluArgThrArgGluArgLeuValTyrGluVal 260

DB 1594 AAACCTAAGGTCAAGAGATCATTTGATGAGCCACAAGGAGCGGCTGTGATGAGGTG 1653

QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280

DB 1654 CCCCAAGAGTGAAGAAATCCAGAGCATTTATCTCATGTGGGAGCTCAATGTCAACC 1713

QY 281 LeuGluHisProLeuPheIleGlyLysMetCysGlnAsnCysLysAsnCysPheLeuGlu 300

DB 1714 CTGAGACACCACTCTTCACTTGAAGCATGTGCCAGAACTGTAAGAACTGCTTGGAG 1773

QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320

DB 1774 TGTGCTTACCAATTAGACAGCATGGGTACAGTCTTATTCACCACTGTGCTGTGGGGG 1833

QY 321 ArgGluValLeuMetCysGlyLysAsnAsnAsnCysArgCysPheCysValGluCysVal 340

DB 1834 CCGTGAAGTCTCATGTGTGGAAACAACACTGCTGAGGTGCTTTGTGTGAGTGTGTG 1893

```

QY 341 AAspleuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTTPaAncyS 360
DB 1894 GATCTCTGTGGGGCCAGAGAGCTGCTCAGAGCAGCATTAAGAAAGACCCCTGAACTGG 1953
QY 361 TTYMeCyGlyIshLysGlyThrTYrGlyLeuLeuArgArgGluAspTTPProSer 380
DB 1954 TACATGTGGGGGCAATAGGGACCTATGGGCTCTGCGAAGAGCGGAGACTGGCTTCT 2013
QY 381 ArgLeuGlnMeCPhePheAlaAenAenH1AspGlnGluPheAspProPolyValTYr 400
DB 2014 CAGCTCAGATGTTCTTTGGCAATTAACATGACAGGAATTTGACCCCAAGGTTTAC 2073
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 2074 CCACCTGTGCGACCTCAGAGAGAAAGCCATCCGGCTGCTGCTCTCTTGAATGGATT 2133
QY 421 AlaThGlyLeuLeuValLeuLysAspLeuGlyIleGlnAlaAspArgTYrIleAlaSer 440
DB 2134 GCTACAGGGCTCTGTGCTGAAGAGCTGGGCACTCAAGTGAACCGCTACATTGCTCC 2193
QY 441 GluValCyGlyAspSerIleThrValGlyMetValArgH1sgnGlyLysIleMetTYr 460
DB 2194 GAGGTGTGGAGGACCTCCATCAGGTGGGCATGTGGGCACAGGAAAGATCATGTAC 2253
QY 461 ValGlyAspValArgSerValThrGlnLysH1IleGlnGluTrpGlyProPheAspLeu 480
DB 2254 GTGGGGGAGCTCCGAGCGTCACAGAGACATATCCAGAGTGGGGCCATTGACCTG 2313
QY 481 ValIleGlySerProCyAspAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 2314 GTGATTGGAGGAGCTCCCTGCATGACCTTCCATTGTCAACCTGCGCGCAAGGACTT 2373
QY 501 TYrGluGlyThrGlyArgLeuPheGluPheTYrArgLeuLysH1sAspAlaArgPro 520
DB 2374 TATGAGGATAGTGGCGGCTCTTGTGATGTACCGCTCTGATGATGCGCGGCCC 2433
QY 521 LysGluGlyAspAspArgProPhePheThrPheGluLysAsnAlaIleMetGlyVal 540
DB 2434 AAGGAGGAGATGATGCGCCCTTCTTGTGCTCTTGAAGATGTGGGCGCATGGGGGTT 2493
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
DB 2494 AGTGACAGAGGACATCTCGCATTTCTTGAAGTCAACCCGATGATGACGCCAA 2553
QY 561 GluValSerAlaAlaH1sArgAlaArgTYrPheTrpGlyAsnLeuProGlyMetAsnArg 580
DB 2554 GAAGTGTGCTGCTCACACAGGCGCCGTTACTTGGGGTAACTTCTGCGCATGAACAG 2613
QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluLysLeuGluH1sgLArg 600
DB 2614 CCTTTGGCATCCACTGTGAATGATGAAGCTGAGACTGCAGAGTGTGAGACGCGACA 2673
QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
DB 2674 ATAGCCAGATTCGCAAAAGTAGAGACCATTAACACAGGCTCAACTTAAGAGAGGCG 2733
QY 621 LysAspGlnH1sAspProValPheMetAsnGluLysGluAspIleLeuTYrCystrGlu 640
DB 2734 AAAGACACAGATTTCCCGCTTCATGAACGAGAGAGAGACATCTCGTGTGCTACAGAA 2793
QY 641 MetGluArgValPheGlyPheProValH1sTYrThrAspAlaSerAsnMetSerArgLeu 660
DB 2794 ATGGAAGAGGTGTTGGCTTCCCGTCCATCAACAGACTCTCCAAACATGAGCGCGTTG 2853
QY 661 AlaArgGlnArgLeuGluArgSerTrpSerValProValIleArgH1sLeuPheAla 680
DB 2854 GCGAGCAGAGATGCTCGGGCCGATCTGAGCGTGCAGCTCCGCACTCTTGCT 2913
QY 681 ProLeuLysGluTYrPheAlaCyVal 689
DB 2914 CCGCTGAAGGATATTTTCTGTGTG 2940

```

```

AY271299
LOCUS AY271299 2597 bp mRNA linear MAM 05-SEP-2003
DEFINITION Bos taurus DNA methyltransferase 3a isoform 4 (DNMT3A) mRNA,
partial cds.
ACCESSION AY271299
VERSION AY271299.1 GI:32350982
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Bovora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 2597)
AUTHORS Golding,M.C. and Westhusin,M.E.
TITLE Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and
Expression in Bovine Preimplantation Embryos, Fetal and Adult
Tissues
JOURNAL Gene Expr. Patterns 3 (5), 551-558 (2003)
REFERENCE 2 (bases 1 to 2597)
AUTHORS Golding,M.C. and Westhusin,M.E.
TITLE Direct Submission
JOURNAL Submitted (08-Apr-2003) Veterinary Physiology, Texas A&M - CVM, 500
University Dr. West, College Station, TX 77843-4466, USA
FEATURES
source
1..2597
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
<1..2597
<1..2597
<1..2467
/gene="DNMT3A"
/contig="DNMT3A"
/feature="DNMT3A4: DNA (cytosine 5) methyltransferase"
/codon_start=2
/product="DNA methyltransferase 3a isoform 4"
/protein_id="AAP75902.1"
/db_xref="GI:32350983"
/translation="EGEEDEBARAEERPESTARKVGRPKRKHPVESSDPRKD
PAVTSKLSMADSGPSLEYPNGDLEKSEEPREGSGAGGKGAPARGEGATEPPE
EASRAVENGCCTPKPEPLQKRAKKAIVIAVMNAVENQSTSSQKYEASPAVOPT
DASAPVTATTPPVADADGNATYAADDEPEYEDGRGFIELVWGLRGRFSWMEK
IVSWMTGRSRAEBGTIRVMWFGDFKFSVCEVKMLSSFCSAFHQATYNKQPMFK
ALYEVLQVASSRAGLFPWCHDSDESDTPAKAVONKQWIEVALGGFQSGRGLPEP
EEKRNPKEVYDMDVMEPEAAAYAPPAPKPKRSKSTTEKPKYKEIIDEETRRRLYEV
ROKCNIDICISGSLNVTLEHPLFEGMCQNCNCPLECAVOYDDDOYQSTICG
GKREYLMCGNNCCGCFVECYDLYVGRPAADAIKEDPMNVCYMGCKTGYLLRRD
DMPSLQMFANHHQDEDPPEPVYPPVPEKRPKIVLSLFGIATGLVLKDLGLOV
DRIYASVECEDSITVGMVNRHQKIMVYGVRSVTOKHIOEWGPFDLVIGSGPNDLSI
VNPARKGLYEGGRGLFFEFYRLHDAIPREGSDRPFMLFEVNVVIVGSDKRIISFL
ENPVMIDAKKEYSAHARAYFMWNLPGMNRPLASTVNDLLEOECLEHRIAKFSKVR
TTTTSNSIKQKDDHFPVFNMEKEDILMCTEMERVFGFPVHYTDSNMSRLARSL
GRSWSVPVIRHLFAPLKEFYACV"
ORIGIN
Alignment Scores:
Score: 0 Length: 2597
Pred. No.: 3716.00 Matches: 680
Score: 99.1% Conservative: 3
Percent Similarity: 98.7% Mismatches: 6
Best Local Similarity: 98.9% Indels: 0
Query Match: 14 Gaps: 0
US-10-623-813-85 (1-689) x AY271299 (1-2597)
QY 1 MetAsnAlaValGlnGluAsnGlnAlaSerGlyLysGlnLysValGluGluAlaSer 20
DB 398 ATGAATGCTGTGAAGAAACAGAGGGGTCACCGAGTCTCAGAAAGGTGGAGGGCCAGT 457
QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 458 CTTCCCGGTGTGAGAGCCAGCCGACCTGGGTCTCCCACTGAGTGCACAGCGCTGAG 517
QY 41 ProValGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60

```

Db 518 CCGGTGGGCGCGATGCGGGGGAAGAAATGCCAAGACCTGACATGATGAAATCCGGAG 577
 Qy 61 TGTGUAAPG1YARG1YPHEG1YIEG1YGLUVAL1TRP1YLYSEUARG1YPHE 80
 Db 578 TACGAGACGGCGGGGCTTTGGCATTTGGGAGCTGGTGTGGGGAATCGCGGGCTTC 637
 Qy 81 SETPTTPROG1YARG1YVAL1SETPTTPMETHTG1YARGSERARG1AALAGLU 100
 Db 638 TCCGTGTGGCCAGCGCGCATTTGTGTGTGATGACGGGCGGAGCCGAGAGGGGAA 697
 Qy 101 G1YTHARG1YVAL1METTRPHEG1YASPG1YLYSPHESERVAL1CYVAL1GLU1Y 120
 Db 698 GGCACCTGTTGGGCTCATGTGGTTCCGAGACGCGAAGTTCTCATGTGTGTGTGGAGAG 757
 Qy 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTYRAsnLYSG1NPRO 140
 Db 758 CTGATGCCCTGAGCTCTTCTGACAGTGTTCACACGAGCCACCTACAAACAGCAGCC 817
 Qy 141 METTYARGLYAR1ALETYRG1UVAL1LEUG1NVAL1ASESERARG1AAG1YLYBLEU 160
 Db 818 ATGTACCGCAGAGGCTTCTACGAGGTCCTGAGTGGCCAGCAGTGGAGGGGAAAGCTG 877
 Qy 161 PHEPRO1ACYHISASP1YASPG1USERASP1YLYSAL1VAL1GLU1YLYS1Y 180
 Db 878 TTTCCATGTGTCATATACACGACGAGAGCGACACTGCCAAGCCCTGGAGGTGCGAGAC 937
 Qy 181 LYSGLIMET1LEGLUTRPA1ALEUG1YGLYPHEG1NPROSEGLYPROLYSG1LYEUG1 200
 Db 938 AACAACAAATGATGAGTGGGCGCTGGAGGGTTCAGCCCTCTGCCCCAAGGGCGCTGAG 997
 Qy 201 PROPTROGLUG1ULYSANSP1YLYSG1UVAL1TYRTHRASP1METTRPA1GLUPRO 220
 Db 998 CCCCCGAAAGAGAGAAAGAACCCCTACAAAGATTACACAGACTGTGGTGTGAACCC 1057
 Qy 221 G1UAL1A1A1A1TYR1A1APROPTROPTROPTRO1A1LYS1YSPROARG1YSETRTHRG1 240
 Db 1058 GAGGACGCTGCTCATGCGCGCCGCCACGACCAAAAGCCCGGAGAGACACAGAGAG 1117
 Qy 241 LYSPTROLYVAL1YSG1UL1LE1LEASP1YARG1YARG1YARG1YARG1YARG1Y 260
 Db 1118 AAGCCTAAGGTCAAGAGATCATTTGATGAACGACAAAGAGGGCGGTGTGCGAGTA 1177
 Qy 261 ARGGLNLYSEYARGHANS1LEG1UASP1YLYS1YSEYSG1YSEYUENVAL1THR 280
 Db 1178 CGCGAAGAGTGGCGGAAACATCGAGGACATCTGCATCTCTTGGGAGCCCTCAACGTCACC 1237
 Qy 281 LEUG1NHI1SPROLEUHE1LEG1YGLYMETCYSG1NAN1CYLSYSAN1CYSPHELEUG1 300
 Db 1238 TTGGAAACACCTCTCTTCATCGAGGAATGTGCCAAACTGCAAGAACTGCTTCTGGAA 1297
 Qy 301 CYBALATYRG1N1YASPA1SPG1YTYRG1N1SETYR1CYSTHR1YLYS1YSG1YGLY 320
 Db 1298 TCGGCTTACAGTACATGACATGATGCTGCTACCTCTGACACATCTGCTCGGGGGG 1357
 Qy 321 ARGGLUVAL1LEU1METCYSG1YAS1NAN1CYCYSEYARG1YSPHESYVAL1GLUCYVAL 340
 Db 1358 CCGGAGGTGCTCATGTGTGGGAAACAATTCCTGACGAGGCTTTGGCGGTGATGTGTG 1417
 Qy 341 ASPL1LEU1VAL1GLYPROG1YAL1AG1N1A1A11LE1YSG1UASP1PROTPAS1CYV 360
 Db 1418 GATCTCTTGTGGTGGGCGAGGGGCGGCGAGCGCATCAAAAGAAAGCCCTGAAACTGC 1477
 Qy 361 TYRMECYSG1YH1S1LYSG1YTHR1YRG1YLEU1EARG1YARG1YARG1YARG1Y 380
 Db 1478 TACATGTGGGGGACAAAGGCACTACGCGCTGCTGGGGGCGGAGCGAGCTGCGCTGT 1537
 Qy 381 ARGLEUG1METPHEPHEAL1AAS1NAN1H1SASG1NGLUPHASP1PROPTROLYSVAL1TYR 400
 Db 1538 CGGCTTCCAGATGTCTTCGCAACAACCAAGACAGGAATTCATCTCCGAAAGGTTTAC 1597
 Qy 401 PROPTROVAL1PRO1AG1YUS1YARG1YPRO1LEARG1YAL1EUS1YLEUPHESP1YLY 420

Db 1598 CCACCTGTCCGACGTGAGAAAGAAAGCCCATCCGGGTGTGTCTCTATTCATGAAAT 1657
 Qy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 Db 1658 GCTACAGGGCTTCTGTGTGTGAAGAGCTTGGGATTCAGGTGAGACGCTACATCGCTCC 1717
 Qy 441 GluValCysGluAspSerIleThrVal1GlyMetValArgHisGlnGlyIleTyr 460
 Db 1718 GAGGTGTGAGAGACTTCCATCAGGTGGGCA1GTGTGGGACCAAGGAAAGATCATGTAC 1777
 Qy 461 ValGlyAspValArgSerValThrGlyNHisIleGlnGlyTrpGlyProPheAspLeu 480
 Db 1778 GTGGGGGACGTCCGAGCGCTTACACAGAGCAATATCCAGAAATGAGGGCCGCTGATCTG 1837
 Qy 481 ValIleGlyIleSerProCysAsnAspLeuSerIleValAspProAlaArgLYSG1YLEU 500
 Db 1838 GTGATTTGGGGGAGTCCCTGCAATGATCTCTCATGTCAACCTTCGCGAAGGAGCTC 1897
 Qy 501 TYRGLUG1YTHRG1YARGLEUPHEG1UPHE1YARGLEUENH1SASP1AARGPRO 520
 Db 1898 TACAGGGGAC1TGCGCGGCTCTTCTTGAGTTACCGCTCTCGCATGATGGCGGCCC 1957
 Qy 521 LYSGLUG1YASPA1SPARGPROPHETRPLeuPheGluAsnValAlaMetGlyVal 540
 Db 1958 AAGAGGAGATGACCGCCCTTCTGCTGTGAGAAATGTGGTGTGCTCGGGCGCTT 2017
 Qy 541 SerAspLYARGASP1LESERARGPHELEUG1USERASP1VAL1MET1LEASP1ALYS 560
 Db 2018 AGTGACAAAGGAGACATCTCGCATTTCTGAGTCCAACTGTGTGATGATGATGCCAAA 2077
 Qy 561 GluValSerAlaAlaHisArgAlaArgTYRPHETRPGLYAS1NLEUPROG1YMETAS1ARG 580
 Db 2078 GAAGTCTCAGCTGCGGACAGGGCCCGCTACTTCTGGGGAACTTCTGTGTATGAACAGG 2137
 Qy 581 PROLEU1ASER1YTHVAL1ASN1LYS1LEUG1YGLUCYSG1EUG1NHI1SG1YARG 600
 Db 2138 CCATTTGCATCCACTGTGATGATGAAGCTGAGCTCAGAGAGTGTGAGAGCGGCCGA 2197
 Qy 601 IleAlaLYSPHESERLYSVAL1ARGTHR1IETRTHRARGSERHNSER1ILEYSG1NGLY 620
 Db 2198 ATAGCCAACTTACGCAAAATGAGAACATTACTACTAGTCGAATCCCAAAAGCAGGGC 2257
 Qy 621 LYSASPGLNHI1SPHEPROVAL1PHEMETANGLULYSGLUASP1LEU1TRPYCYTHRG1 640
 Db 2258 AAGAGCACACATTTCCCGCTTTCATGATGAGAAAGAGACATCTTATGTGTACATGAA 2317
 Qy 641 METGLUARGVAL1PHEG1YSPHEPROVAL1H1TYRTHRASP1SERAN1SETSERARGLEU 660
 Db 2318 ATGGAAGGTTGTGGTTCCTTCCTGTCACTATACCGAGGTCTCCAAACATGAGCCGCTTG 2377
 Qy 661 AlaArgGlnArgLeuLeuGlyArgSerTRP1SERVAL1PROVAL1LEARGH1SLEUPHEAL1A 680
 Db 2378 GCGAGGCAAGAGACTGTGGCGGCTGTGTGAGAGGTGCGCGGTATCCGCCACTTTCGCT 2437
 Qy 681 PROLEU1YSG1UTYRPH1A1CYVAL 689
 Db 2438 CCGCTGAAGGAATATTTCCTGTGTG 2464

RESULT 5
 AY271298 2798 bp mRNA linear MAM 05-SEP-2003
 LOCUS Bos taurus DNA methyltransferase 3a (DNMT3A) mRNA, partial cds.
 DEFINITION
 AY271298
 ACCESSION
 VERSION AY271298.1 GI:32350980
 KEYWORDS
 SOURCE Bos taurus (cattle)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 REFERENCE
 1 (bases 1 to 2798)
 1 Golding M.C. and Westhusin M.E.
 ANALYSIS of DNA (cytosine 5) Methyltransferase mRNA Sequence and

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Expression in Bovine Preimplantation Embryos, Fetal and Adult
Tissues
Gene Expr. Patterns 3 (5), 551-558 (2003)
2 (bases 1 to 2798)
Golding, M.C. and Westhusin, M.E.
Direct Submission
Submitted (08-APR-2003) Veterinary Physiology, Texas A&M - CVM, 500
University Dr. West, College Station, TX 77843-4466, USA

FEATURES

source
1. 2798
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
<1. 2798
/gene="DNMT3A"
<1. 2668
/note="DNMT3A"
/codon_start=2
/product="DNA methyltransferase 3a"
/protein_id="AAP75901.1"
/translation="EGEBOBARAKEROBPSTTARKVGRGKRKHPPVSSDTPKD
PVTSTSLMAODSPGSELYPNGDLKRSPOBEGSPAGGKGAPABEGATERP
ENSRAYENGCCTKDGRCAPAEBSKEQKINISMCEGSRGLRGGLMESLRRP
MPLRTFOAGDPYITSKRKDEMLARKRELEKAKYIAVWNAVEENOGSTESQKEA
SPAVOQPDPAPEVATVTPPEVADAGDNATKADDEBEDGSGFGLVGLWKL
RGSWMGRIVSWMTGRSRAAEGTRVMVFGDKFSVVCVECLMSEFSCAFHQT
YNKQPMYRAKIVLQVASSRAKLPFMCCHSDSDPTAKAEVONKQMLMALGQOP
SGPKGLEPBEENPYKEVYTDWVPEAAVAPPAKPRKSTKTEPKVXKIIIDER
TERRLYEVRQKCRNIEDICISGSLNLTLEHPLFCGMCONCKNLEFCAYDDG
YOSYCTCCGREGVLMCGNNCCRCFCEVCLLVGRQAQAIAKDPNMCYCGKRG
TYGLARRDMPRLQMFANNDQEPDPKVPVPAEKRIKVLSDGLATLIL
VXKDLGIVQDRIYASEVCEDSITVGMWRHQKIMYDVNSVYQKIQEMEPDLVIG
CEGNDLSIVNPARKGLYEGTGRLEFEFYLLDAPKEDDPPFMLENVLVGS
DNRDISRLPSNPMVIMDAKESAAHARVYMGNLPGMNPRLASTVDKLEOCLBHG
RIAKESKVRITITRSNISKQKDOHPVFNENEKILMCTEMRVRGFPVHYDVSM
SRLAROLIGRSMVPIRHLPLKEFYFACV"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 2798
Score: 3716.00 Matches: 680
Percent Similarity: 99.1% Conservative: 3
Best Local Similarity: 98.7% Mismatches: 6
Query Match: 98.9% Indels: 0
DB: 14 Gaps: 0

US-10-623-813-85 (1-689) x AY271298 (1-2798)

QY 1 MetAsnaIaValGluGuaEngIaAaSerGlyGluSerGlnLysValGluGluAaSer 20
DB 599 ATGAATGCTGTGAAGAAACCAAGGGGCTCCACCGAGCTCAGAAAGTGGAGGCGCAGT 658
QY 21 ProProAlaIaValGlnGlnProThrAspProAlaSerProThrValAlaIaThrTrpProGlu 40
DB 659 CCTCCCGCTGTGAAGCCACCGACCTGCGCTCCGCAAGTGGACCAAGCCTGAG 718
QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaIaAaAspAspGluProGlu 60
DB 719 CCGGTGGGCCCGATGCGGGGCAAGAAATGCCACCAAGCAGTGCATGAAACCGGAG 778
QY 61 TyrGluAspGlyArgGlyPheGlyTlGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
DB 779 TACGAGACGCGCGGCTTTGGCATTTGGGAGAGCTGTGTGGGGAAACTGCGGGCTTC 838
QY 81 SerTrpTrpProGlyArgGlyLeuAlaSerTrpTrpMetThrGlyArgSerArgAlaIaGlu 100
DB 839 TCTGTGGGCGAAGCCCATTTGTCTTGTGGATGACGGGGCCGAGACCGACGCGGAA 898
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB 899 GGCACCTGTTGGGTCAATGTGTTGCGAGACGGCAAGTTCTCAGTGTGTGTGAGAAAG 958

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheIleGlnAlaThrTyrAsnLysGlnPro 140
DB 959 CTGATGCGCGTAGAGTCTCTTGTGCAGTGTCTTCCACAGGCGCACTTACACAAAGCAGCCC 1018
QY 141 MetTyrArgLysAlaIleTyrGluValIleuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 1019 ATGTACCGAAGGCTATCTACAGAGGCTCTGACAGTGGCCAGCAGTCCAGCGGGAAGCTG 1078
QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAan 180
DB 1079 TTTCAATGTGCATGACACGACACAGCAGCAGCTGCGCAGGCGGTGAGGTGCGCAAC 1138
QY 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 1139 AAACAAATGATGAGTGGGCTCTGGAGAGGTTCCAGGCTCTGCGCCCAAGGCGCTCGAG 1198
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
DB 1199 CCCCAGAAAGAGAGAGAACCCCTTACAAAGAGTTACACAGACATGTGGTTGAACCC 1258
QY 221 GluAlaAlaIleTyrAlaProProProProAlaLysLysProArgLysSerThrTrpGlu 240
DB 1259 GAGGACAGCTGCTTATGCGCGCCGCCACGACCAAAAGCCCCGAAAGACCAACGAG 1318
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLysValTyrGluVal 260
DB 1319 AAGCTTAAGGTCAAGAGATCAATTGATGAACGCAAGAGAGCGCTGTTACAGATN 1378
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1379 CGCGAAGATGCGGGAATCGAGGACATCTCATCTTGTGGAGCTTCAACGTCACCC 1438
QY 281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 1439 TTGGAAACCCCTCTCTTCTCGAAGAAATGTCACAAATCGCAAGATGCTTCTTGGA 1498
QY 301 CysAlaIaTyrGlyTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB 1499 TGCGCTACCATGATCATGATCATGCTCATGCTTATGCAACATCTGCGGGGG 1558
QY 321 ArgGluValIleuMetCysGlyAsnAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
DB 1559 CGGAGAGTCTCATGTGTGGAAACAATTCGTGCGAGTGTCTTTCGTGGAATGTGTG 1618
QY 341 AspLeuLeuValGlyProGlyValAlaGlnAlaIleLysGluAspProTyrAsnCys 360
DB 1619 GATCTCTTGTGGGCGCAGGGCGCGCGAGGACCATCAAGAAAGACCCCTGGAACCTGC 1678
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGlyLysAspTrpSer 380
DB 1679 TACATGTGGCGGACAGGACGACCTTACGGGCTGTGCGGGGGGAGCAGCTGGCGTCT 1738
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProlysValTyr 400
DB 1739 CGGCTCAGATGTTCTTCCCAACAACATGACACAGAAATTCATCTCTCGAAGGTTTAC 1798
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 1799 CCACTGTCTCCAGCTGAGAAAGAAAGCCATCTGGGTGTCTTCTTATTCATGGAATT 1858
QY 421 AlaThrGlyLeuLeuValIleuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
DB 1859 GCTACAGGGCTTGTGTGCTGAAGACTTGGGCAATTCAGGTGACCGCTATCATGCTCC 1918
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlnLysIleMetTyr 460
DB 1919 GAGGTGTGAGGACCTCATCAAGTGGCATGTGTGGCAGCAAGGAAAGATCATGTAC 1978
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
DB 1979 GTCGGAGAGCTCGCAGCGCTTACACAAAGCATATTCAGGAATGGGCGCGTTCGATCTG 2038

Qy	48	Val111eg11g1SePrOcYsaAmSpleuSer111eValAsPrOAlArGySe111eU	500
Db	2039	GTGATTGGGGGCAAGTCCTCCGCAATGATCTCTCACTGTCACACCTGCCGCAAGGACTC	203989
Qy	501	TYRG1UG1YThG1YArG1eUhePheG1uPheTYrArG1eUheH1AsPa1AaRPro	520
Db	2099	TACGAGGGCACTGGCCGGCTTCTTGTAGTTCTACCGGCTCTCATGATGACCGGCCC	2158
Qy	521	LYSG1UG1YAsPaARgProPhePheTTPLeUPheG1UAsnVal1Val1AmeC1YVal	540
Db	2159	AAGGAGGGAGATGACCGCCCTTCTTGGCTTGTGAAATGTGTGTCTGTGGCGCTT	2218
Qy	541	SeRaSPlYArGAsPl1eSeRaRPhLeUeG1USeRaSnProValMe11eAsPa1AaYs	560
Db	2219	AGTGACACAGGAGCAATCTCGCGATTTCTCGAGTCCAACTGTATGATTTGATGCCAA	2278
Qy	561	GIUVal1SeRa1AAlAH1sARg1AaRgTYrPheTTPG1YAsnLeUProG1YMeCAsnARg	580
Db	2279	GAAATGTACGCTGGCCACAGGGCCCGCTACTTCTGGGGGAACCTTCCTGTATGAACAGG	2338
Qy	581	ProLeUAlAsErThValAsnAsPlYsLeUeG1UeG1UeG1UeG1UeG1UeG1UeG1YArG	600
Db	2339	CCATTGGCATCCACTGTGATGATGAAGCTGGAGCTGCAGAGTGTCTGAGACACGGCCGA	2398
Qy	601	11eAl1AlYsPheSeR1YsValArGrh1r1eThTTPArGSeRaSnSer111eLYeG1UG1Y	620
Db	2399	ATAGCAAGTTCAGCAAAAGTAGAGCACTACTACTAGTGCAGACTCATTAACAGAGGC	2458
Qy	621	LYsARPG1NH1sPheProValPhemeRAsnG1UySG1UAsPl1eLeUTPcYSThG1U	640
Db	2459	AAAGCCAGCATTTCCCGCTCTCATGAATGAAAGAGGACATTTATGTGTGACTGAA	2518
Qy	641	MeCg1UArGVal1PheG1YpHeProValH1sTYrThAsPa1SeRaSnMeSeRaRLeU	660
Db	2519	ATGGAAGGGGTGTGGCTTCCCTGTCCACTATACCGACGTCCCAACATGACGGCTTG	2578
Qy	661	Al1ARg1NArG1eUeUeG1YArGSeRTPSeRaValProVal11eArG1H1sLeUPheAla	680
Db	2579	GCGAGGCGAACACTGTGGCGCGGTCTGTGAGCGTCCGATCATCCGCACTTCTTGCTT	2638
Qy	681	ProLeUySG1UTYrPheAlaCYsVal	689
Db	2639	CCGCTGAAGGAATATTTTGTCTTGTGTG	2665
RESULT 6			
AF480163		2371 bp	mRNA
LOCUS			linear PRI 14-OCT-2002
DEFINITION			Homo sapiens DNA cytosine methyltransferase 3a2 (DNMT3A2) mRNA,
ACCESSION			complete cds.
VERSION			AF480163
KEYWORDS			AF480163.1 GI:23954437
SOURCE			.
ORGANISM			Homo sapiens (human)
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
			Homnidae; Homo.
			1 (bases 1 to 2371)
REFERENCE			Chen,T., Ueda,Y., Xie,S. and Li,E.
AUTHORS			A Novel Dmrt1 Isoform Produced from an Alternative Promoter
TITLE			Localized to Euchromatin and Its Expression Correlates with Active
			of Novo Methylation
JOURNAL			J. Biol. Chem. 277 (41), 38746-38754 (2002)
PUBMED			12138111
REFERENCE			2 (bases 1 to 2371)
AUTHORS			Chen,T. and Li,E.
TITLE			Direct Submision
JOURNAL			Submitted (31-JAN-2002) Cardiovascular Research Center,
			Massachusetts General Hospital, 149 13th Street, Charlestown, MA
			02129, USA
FEATURES			Location/Qualifiers
SOURCE			1..2371
			/organism="Homo sapiens"

gene
CDS

1. .2371
/gene="DNMT3A2"
217. .2286
/function="de novo DNA cytosine methylation"
/codon_start=1
/product="DNA cytosine methyltransferase 3A2"
/protein_id="AA04037.1"
/db_xref="GI:23954438"
/translation="MNANIEBNOGREGESQKVEASPPAVQOPTHDPASPIVATTPREYVG
DAGDNATKAGDEPEYEDGRGFGIGELVWGELKRFSSWPGRIIVSMWMTGSRRAEG
RWVMVFSGDGKRSVVCVEKLMPLEISFCSAFPHATYKNKQMPYRAIYEVQVASSRQAK
FVCKDESDSDTLAKAVEVONKPMIEALGFQSPGKLEPPEEKNPKEYVTDMM
EPEAAAYAPPPAKKPKRSTAEKPKVKEIIIDERTBERLVYEVORCRNIEDI CISGCG
LVNTEHPLFVCGMCONKCNKPLFCAYODDDGYOSYCTTCGGEVLMCONNCR
FCVECVDLIVPGAAQAAIKEDPWNCTWGHKTYTGILRRREDPRLQMFPPANNHD
EEDPKYVPVPAERKRPKIRVLSPFDLGIATLILVKDIGIYDRIYASEVEDSITTV
MVRHGKIMYGVDSVYQKHIQEWPGFDLVIIGSGPCNDLSIVNPARGLVEGTORL
FEFYRLHDARPKEGDDPFPMLFENVAVMSVDRDLSRFLSNPVMIDAEVAAAL
RABRYWNLPGMARNPIASTVNDKLEOCELEHGRARFSKRTITTRENSIIKQGDND
FPPVNEKEDILNCTEMERVRFPHVITDYVSNMRSLARQLLRGMSVPIVRLHLPAP
KEYPACV"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2371
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatch: 8
Query Match: 98.8% Indels: 0
DB: 5 Gaps: 0

US-10-623-813-85 (1-689) x AF480163 (1-2371)

QY	1	MeCaenAaValaGluGluAaSnGlnAlaSerGlyGluSerGlnIySvaGluGlnAlaSer	20
Db	217	ATGAATGCTGTGGAAAGAAAACCAAGGGGCCGGGGAGTCTCAGAAAGTGGAAGAGCCACG	276
QY	21	ProPfoAlaValaGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu	40
Db	277	CCTCGTGTGTGGACAGCCACCACTGACCCGCCACCTCCACTGTGGGTACACACGCTTAG	336
QY	41	ProValaGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu	60
Db	337	CCCGTGGGGTCCGATGCTGGGGGACAAGATGCCACCAAGACGGCGAAGACGACGACGACG	396
QY	61	TyrGluAspGlyAraGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuAraGlyPhe	80
Db	397	TACAGAGACGGCCGGGGCTTTGGCATTTGGGAGCTGTGTGTGGGGGAAACTCGGGGCTTC	456
QY	81	SerTyrTyrProGlyAraGlyLeuAlaSerTyrTyrMetThrGlyAraGlySerAraGlyAlaGlu	100
Db	457	TCTCGTGGCCAGAGCCGCAATGTGTCTTGTGTGATGACGGGGCCGAGCCGACGAGCTGAA	516
QY	101	GlyIhrArgTyrPylMetTyrPheGlyAspGlyLysPheSerValAlaCysValaGluLys	120
Db	517	GGCACCGCGTGGTGCATGTGTGTGGAGACGGCAAAATTTCTCAGTGTGTGTGGAGAG	576
QY	121	LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro	140
Db	577	CTGATGCCGCTGAGACTGCTTTTGCAGTGCCTTCCACACAGGCCACGTACACAAACACGCC	636
QY	141	MetTyrArgLysAlaIleTyrGluValLeuGlnValaAspSerAraGlyAlaGlyLysLeu	160
Db	637	ATGACCGCAAAAGCCATCTACGAAGCTCTCGAGGTGGCCAGCACCGCGCGGGAAGCTG	696
QY	161	PheThrAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValaGluValaGlnAsn	180
Db	697	TTCCCGGTGTGCCACGACGAGTGAAGTGAACACTCCCAAGGCCGTGTGAGAGTGCAGAAC	756

```

Qy 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGln 200
Db 757 AAGCCATGATTAATGAGGCCCTGGGGGCTTCCAGCCTTGTGGCCCTTAAGGCGCTGAG 816
Qy 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db 817 CCACCAAGAGAGAGAAATCCCTCAACAAAGATGTAACAGGACATGTGGTGAACCT 876
Qy 221 GluAlaIleAlaTyrAlaProProProAlaLysLysProAlaGlySerThrGlu 240
Db 877 GAGGAGCTGCTTACGACCACTCCACCAACCAAAAGCCCGGAAGACAGCCGAG 936
Qy 241 LysProLysValLysGluIleIleAspGluArgThrArgLysLeuValTyrGluVal 260
Db 937 AAGCCCAAGGTCAAGGATTAATGATGAGCGCAACAGAGCGGCTGTGTAAGAGGTG 996
Qy 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 997 CGGCGAAGGTGCGGAACTTGAGGACATCTGCATCTCTGTGGAGCCTCAATGTTACC 1056
Qy 281 LeuGlnHisProLeuPheIleGlyGlyMetCysGlnAsnGlyLysAsnGlyPheLeuGln 300
Db 1057 CTGGAACACCCCTCTTGTGTGGAGAAATGTGCCAAACTGCAAGAACTGCTTCTGAG 1116
Qy 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1117 TGTGCGTACCACTACGACGACGACGCGCTACAGTCTCACTGACCATCTGCTGTGGGGG 1176
Qy 321 ArgGluValLeuMetCysGlyLysAsnAsnGlyCysArgCysPheCysValGluCysVal 340
Db 1177 CGTGAGGTGCTCACTGTGGGAAACAACTGTCAGAGTCTTTGTGGTGAATGTGTG 1236
Qy 341 AspleuLeuValGlyProGlyAlaIleGlnAlaIleLysGluAspProTrpAsnGly 360
Db 1237 GACCTCTTGTGGGGCGGGGGCTGCCACGACCACTTAAGAAAGCCCTGGAATCGC 1296
Qy 361 TyrMetCysGlyHisLysGlyTyrTyrGlyLeuLeuArgArgGluAspTrpProSer 380
Db 1297 TACATGTGCGGGCAACAGGTACTCAACGGCTGTGCGCGGCGAGAGACTGCGCTCC 1356
Qy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProLysValTyr 400
Db 1357 CGGCTCAGAGTCTTCTGCTAATTAACCAACAGCAAGAAATTGACCTCCAAAGGTTTAC 1416
Qy 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 1417 CCACCTGTCCACGCTGAGAAAGAGAGAGCCCATCCGGGTGCTGTCTCTTGTGATGGAATC 1476
Qy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 1477 GCTACAGGGCTCTGTGTCTGTAAGGACTTGGGCAATTCAGTGGACCGCTACATTCCTCG 1536
Qy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 1537 GAGGTGTGTGAGCACTCATCAGGTGGGCATGTGGCGCACAGGGAGAAATCATGTATC 1596
Qy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db 1597 GTCGGGAGCTGCCGAGCTGACACAGAACATATCCAGAGTGGGGCCATTCGATGTG 1656
Qy 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db 1657 GTGATTTGGGGGAGTCTCCCTGCAATGACTCTTCATCTCCCACTGCTCCGCAAGGCTTC 1716
Qy 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 1717 TACAGAGGAGCTGGCGGCTCTTCTTGAATTTCAACCCCTCTGTCATGAGCGGCGCC 1776
Qy 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
Db 1777 AAGAGAGGAGATGATCGCCCTTCTTCTGAGCTCTTGTGAAGATGTGTGGCATGGGCTT 1836

```

```

Qy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 1837 AGTGAACAGAGGAGCACTTCGCGATTTCTCGAATTCACACCTGTGATGATGACCAAA 1896
Qy 561 GluValSerAlaIleHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db 1897 GAAGTGTACGCTCAACACAGGCGCCCTACTCTTGGGGTAACTTCCGGTATGAACAGG 1956
Qy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnLysLeuGlnHisGlyArg 600
Db 1957 CCGTGGCATCCACTGTGAATGATTAAGCTGAGCTCAGAGAGTGTGGAGCATGGCAGG 2016
Qy 601 IleAlaLysHisSerLysValArgThrIleThrArgSerAsnSerIleLysGlnGly 620
Db 2017 ATAGCCAAAGTTCAGCAAGAGAGAGACCATTAACAAGGTCCTCAATTAAGACAGGC 2076
Qy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db 2077 AAGACCAAGATTTTCTGCTTCTGATGAATGAGAAAGAGACATCTTATGTGCATGAA 2136
Qy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2137 ATGGAAGGATATTGATTGTTCCAGTCCACTATTAAGACCTCTCCACATGAGCGCTTG 2196
Qy 661 AlaArgGluArgLeuGluGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2197 GCGAAGCAGAGACTGCTGGCGGCTGCTGAGAGCGTCCCACTCATCCGCCACTTCTGCT 2256
Qy 681 ProLeuLysGluTyrPheAlaCysVal 689
Db 2257 CCGCTGAAGAGATATTTTGTGTGTG 2283

```

```

RESULT 7
CS050243
LOCUS CS050243 3005 bp DNA linear PAT 23-MAR-2005
DEFINITION Sequence 27 from Patent WO2005021757.
ACCESSION CS050243
VERSION CS050243.1 GI:61889498
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 Tomme, P.H. and van Rompaey, L.
TITL Polypeptides and polynucleotides for use as a medicament
JOURN Patent: WO 2005021757-A 27 10-MAR-2005;
Galapagos Genomics N.V. (BB)
FEATURES
source location/Qualifiers
1..3005
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3005
Score: 3714.00 Matches: 679
Best Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
DB: 2 Gaps: 0
US-10-623-813-85 (1-689) x CS050243 (1-3005)
Qy 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyLysSerGlnLysValGluGluAlaSer 20
Db 899 ATGAATCTGTGAAGAAACAGAGGGCGCGGAGAGTCTCAAGAGGTGAGAGAGCCAGC 958
Qy 21 ProProAlaValGlnGlnProTrpAspProAlaSerProThrValAlaThrTrpGlu 40
Db 959 CTTCTCTGTGAGCAGCCCACTGACCCCGCATCCCTCCACTGTGTGCTACACCGCTGAG 1018

```

QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 Db 1019 CCCGGGGGCTCCGATGCTGGGGACAAGATCCACCAAGCGCGGATGACGAGCCAAAG 1078
 QY 61 TyGluAspGlyArgGlyPheGlyY1IeGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
 Db 1079 TACAGAGACGGCCGGGGCTTTGGCATTGGGAGCTGGTGGGGAAATCCGGGGCTTC 1138
 QY 81 SerTrpTrpProGlyArgGlyLeuAlaSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 Db 1139 TCCCGTGGCCAGCGCCGATGTGTCTTGATGAGCGGGCCGAGCGGAGAGCTGAA 1198
 QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 Db 1199 GGCACCCGCTGGTCAATGTGTGGAGACGGCAAAATTCAGTGTGTGTGTGAGAG 1258
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTrpAsnLysGlnPro 140
 Db 1259 CTGATGCCGCTGAGCTCCTTTTGCAGTGCCTTCCACGAGCCAGCTACAGACAGCC 1318
 QY 141 MetTrpArgLysAlaAlaTrpGlyValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 Db 1319 ATGACCGCAAGCCATCTACGAGGTCTCGACGGTGGCCAGACCGCGCGGAGAGCTG 1378
 QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValAsn 180
 Db 1379 TTCCGGGTGGTCCACGACGAGATGAGTACATCTCCAGGCCGTGGAGCTCAGAAC 1438
 QY 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProGlyLeuGlu 200
 Db 1439 AAGCCCATGATTGAATGGCCCTGGGGGCTTCCAGCTTCTGGCCCTTAAGGGCTGGAG 1498
 QY 201 ProProGluGluGlyLysAsnProTrpLysGluValTrpThrAspMetTrpValGluPro 220
 Db 1499 CCACCAAGAGAGAGAAATCCCTCAAAAGATGTACAGGACATGTGGTGAACCT 1558
 QY 221 GluAlaAlaAlaAlaTrpAlaProProProAlaLysLysProArgLysSerThrThGlu 240
 Db 1559 GAGGACCTGCTTACGACCCCTCCACAGCCAAAGAGCCCGGAGAGAGACAGCGAG 1618
 QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTrpGluVal 260
 Db 1619 AAGCCCAAGGTCAAGAGATTTATGATAGCGCACAGAGCGGCTGTACAGAGTG 1678
 QY 261 ArgGluLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 Db 1679 CGGCAAGATGCCGGAACATTTGAGACATCTCTGTGGAGGCTCAATGTAAAC 1738
 QY 281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 Db 1739 CTGAAACACCCCTCTTCTGTGGAGAAATGTGCCAAACGCAAGAACTGCTTTCTGGAG 1798
 QY 301 CysAlaTrpGlnTrpAspAspAspGlyTrpGlnSerTrpCysThrIleCysCysGlyGly 320
 Db 1799 TGTGCGTACCGTACGAGAGAGAGCGGCTACCACTCTCTACACATCTGTGTGGGGC 1858
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 Db 1859 CGTGAAGTGTCTCATGTGCGGAAACAACTGTCAGGTCTTTTCTGTGAGTGTG 1918
 QY 341 AspLeuLeuValGlyProGlyValAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
 Db 1919 GACCTTGTGGGGCGCGGGGCTGCGCCAGGACACCTTAAGAAACCTCTGGAACCTG 1978
 QY 361 TyrMetCysGlyHisLysGlyThrTrpGlyLeuLeuArgArgGluAspTrpProSer 380
 Db 1979 TACATGTGGCGGACAAAGGATCTTACGGGCTGTGGCGCGGACAGAGACTGGCCCTCC 2038
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTrp 400
 Db 2039 CGGCTCAGAGTCTTCTGCTAATTAACACAGACAGAAATTTGACCCCTCAAAAGTTTAC 2098

QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 Db 2099 CCACCTTCCCACTGTGAAG 2158
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyY1IeGlnValAlaAspArgTrpIleAlaSer 440
 Db 2159 GCTACAGAGCTCCCTGGTGTGAAGAGACTTGGGGCATTCAGGTGGAGCCCTACATTTGCC 2218
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTrp 460
 Db 2219 GAGGTGTGTAGAGACTTCATCACGGTGGATGTGGGACCAAGGGAGATATATGTAC 2278
 QY 461 ValGlyAspValArgSerValIleGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
 Db 2279 GTCCGGAGCTCCCGACGTCACACAGAGACATTCACAGAGTGGGCCCATTTGATCTG 2338
 QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 Db 2339 GTGATTTGGGGAGTCCCTGCAATGACCTCTCCATCGTCAACCTCTCGCAAGGGCCTC 2398
 QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTrpArgLeuLeuHisAspAlaArgPro 520
 Db 2399 TACGAGGACCTGGCCGCTCTTCTTGAAGTCTTACCGCTCTGCAATGATGCGGGCC 2458
 QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaAlaMetGlyVal 540
 Db 2459 AAGGAGGAGATGATGCCCTCTTCTTGTGGCTCTTGAAGATGTGTGGCCATGGCGTT 2518
 QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
 Db 2519 AGTGACAAAGAGGACATCTCGCATTTCTCGAGTCCAAACCTGTGATGATGAGCCAAA 2578
 QY 561 GluValSerAlaAlaHisArgAlaArgTrpPheTrpGlyLysLeuProGlyMetAspArg 580
 Db 2579 GAAGTGCACCTCCACACAGAGCCCGCTACTTCTGGGTAACTTCCCGATGAACAGG 2638
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnGluCysLeuGluHisGlyArg 600
 Db 2639 CCGTTGGCATCCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2698
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 Db 2699 ATAGCCAACTTACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2758
 QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
 Db 2759 AAAGACCAAGATTTCTGTCTTCAATGAATGAAGAAAGAGAGAGAGAGAGAGAGAG 2818
 QY 641 MetGluArgValPheGlyPheProValHisIleTrpThrAspValSerAsnMetSerArgLeu 660
 Db 2819 ATGGAAGGATATTTGTTTCCAGTCCATATATGACCTCTTCCACATGAGCGGCTTG 2878
 QY 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
 Db 2879 GCGAGGACAGAGACTGCTGGGCGGATCATGAGAGCGTGCAGTCACTCCGCACTTCTGCT 2938
 QY 681 ProLeuLysGluTrpPheAlaCysVal 689
 Db 2939 CCGCTGAAGAGATATTTGTGTGTGTG 2965
 RESULT 8
 AF067972 3005 bp mRNA linear PRI 12-FEB-2001
 LOCUS Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,
 DEFINITION complete cds.
 ACCESSION AF067972
 VERSION AF067972.2 GI:12746531
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

QY 421 AATThrGlyLeuLeuValLeuLysAspLeuGlyIleGluValAspArgTyrIleAlaSer 440
 DB 2159 GCTACAGGGCTCTGCTGCTGGAAGGACTTGGGCACTCAGGTGACCGGCTACATTGCTCTG 2218
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2219 GAGGTGTGTGAGGACTCCATCAGGTGGGCGATGTGTGGGACAGGAGGAGATCATGTAC 2278
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrPheLeuAspLeu 480
 DB 2279 GTGGGGAGAGCTCCGACGCTCACAGAAAGCAATCCAGAGGGGCCCATTCGATCTG 2338
 QY 481 ValIleGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2339 GTGATTGGGGGAGCTCCCTGCAATGACTCTCCATGTGTACACCTGCTGCAAGGGCTTC 2398
 QY 501 TyrGlnGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2399 TACGAGGGCACTGGCCGGCTCTTCTTGAGTTCTACCGCTCCTGATGATGGGGCCC 2458
 QY 521 LysGlnGlyAspAspArgProPhePheTyrPheGluAsnValValAlaMetGlyVal 540
 DB 2459 AAGGAGGAGATGATGCCCCCTTCTGCTGCTTGGAGAAATGTGTGGCCATGGGCGTT 2518
 QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
 DB 2519 AGTACAAAGAGGACATCTCCGATTTCTCGAGTCCACCCCTGTGATGTATGATGCAAA 2578
 QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyLysAsnArg 580
 DB 2579 GAGGTGACGTGCACACAGGGCCCGCTACTTCTGGGGTAACCTCCGGTATGAAAGG 2638
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnGluCysLeuGlnHisGlyArg 600
 DB 2639 CCGTTGGCATCCACTGTGATGATGATGAGCTGAGCTGAGAGTGTCTGGAGCATGGCAGG 2698
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 DB 2699 ATAGCCAGTTCCAGAAAGTGAAGTGAACCTTACTACAGAGTCAACTCCATAAGAGAGGCTC 2758
 QY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGluAspIleLeuTyrCysThrGlu 640
 DB 2759 AAAGCCAGCATTTTCTGCTTCTGATGATGAAAGAGGACATTTATGGTGCATGA 2818
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2819 ATGAAAGGTAATTTGCTTCCAGTCCACTATCTGACGTCTCCAAATGATGCGGCTTG 2878
 QY 661 AlaArgGlnArgLeuLeuGlyArgSerTyrSerValProValIleArgHisLeuPheAla 680
 DB 2879 GCGAGGCGAAGACTGCTGGCGGATCATGAGCGGTGCCAGTATCCGCCACTTTCCT 2938
 QY 681 ProLeuLysGlnTyrPheAlaCysVal 689
 DB 2939 CCGCTGAAGAGTATTTGGCGTGTG 2965
 RESULT 9
 CS050242 3604 bp DNA linear PAT 23-MAR-2005
 LOCUS Sequence 26 from Patent WO2005021757.
 DEFINITION CS050242
 ACCESSION CS050242
 VERSION CS050242.1 GI:61889497
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.
 REFERENCE
 1 Tomme, P. H. and van Rompaey, L.
 Peptides and polynucleotides for use as a medicament
 Patent: WO 2005021757-A 26 10-MAR-2005;
 JOURNAL

FEATURES Galapagos Genomics N.V. (BE)
 source Location/Qualifiers
 1..3604
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3604
 Score: 3714.00 Matches: 679
 Percent Similarity: 98.8% Conservative: 2
 Best Local Similarity: 98.5% Mismatches: 8
 Query Match: 98.8% Indels: 0
 DB: 2 Gaps: 0
 US-10-623-813-85 (1-689) x CS050242 (1-3604)
 QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnLysValGluGluAlaSer 20
 DB 217 ATGAATGCTGTGAGAAACCAAGGGCCCGGAGTCTCAGAAAGTGAAGAGGCGCAC 276
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 DB 277 CCTCTGCTGTGACAGCACCCACTGACCCCGCATCCCATCTGTGTGATCCACGCTGAG 336
 QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 337 CCGGTGGGGTCCGATGCTGGGAGACAAAGATGCCAAAGAGCGATGACGAGCGACAG 396
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
 DB 397 TACGAGAGCGCCGGGCTTTGGCATTTGGAGCTGTGTGGGAAACTGGGGGCTTC 456
 QY 81 SerTyrTyrProGlyArgIleValSerTyrTyrMetThrGlyArgSerArgAlaAlaGlu 100
 DB 457 TCTGTGTGGCAGGCGCGATTTGTCTTGTGTGATGAGCGGCGGAGCGGACACTGA 516
 QY 101 GlyThrArgTyrValMetTyrPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 517 GGCACCGCGTGGTCACTGTGTGCGAGACCGCAATCTCAAGTGTGTGTGAGAG 576
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 DB 577 CTGATGCCCTGACCTGCTTTTGCATGCTTCCACAGGCGAGTCAACAGAGAGCC 636
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 637 ATGTAACGCAAAAGCATCTACAGAGTCTTGAGGTGGCCAGCCGCGGGAGAGCTG 696
 QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValAsn 180
 DB 697 TTCCCGGTGTCACGACAGGATGAGATGACCTGCCAAGGCGCTGTGAGAGTGCAGAAC 756
 QY 181 LysGlnMetIleGluTyrPalaLeuGlyLysPheGlnProSerGlyProLysGlyLeuGlu 200
 DB 757 AAGCCCAATGATGATGATGAGCCCTGGGGGCTTCCAGGCTTCTGACCTTAAAGGCGCTGAG 816
 QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTyrValGluPro 220
 DB 817 CCACCAAGAGAGAGAAATCCCTACAAAGATGACAGGACGATGTGGTGAACCT 876
 QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrGlu 240
 DB 877 GAGGACACTGCTTACGACACCTCCACAGCCCAAAAGCCCGGAGAGAGACGCGAG 936
 QY 241 LysProLysValLysGluIleLysAspGluArgThrArgGluArgLysValTyrGluVal 260
 DB 937 AAGCCCAAGTCAAGAGATTTATGATGAGGCAACAGAGAGCGGCTGTATGAGAGT 996
 QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 997 CCGGAGAAAGTCCGGAACATTTGAGGACATCTGCTGTGGAGCTTCAATGTTAAC 1056

OY	281	LeuGLuHISpRoLeuPheIIeGIyGlyMeTcYseGlnIaMencySlySaMcySphLeuGlu	300
Db	1057	CTGGAAACACCCCTCTTGTTGGAGGAAATGGCCAAACCTGCAAGAACTGTTCTTGAG	1116
OY	301	CysAlaTYrGlnTYrAspAspAspGlyTYrGlnSeTYrCYsThrIIeCYsGlyGly	320
Db	1117	TGTGGCTACCACTACGACGACGCGCTACCACTCTACCACTCTGCTGGGGGGC	1176
OY	321	ArgGluValLeuMeTcYsGlyAsnAsnAsnCYsCySarGcYsPheCYsValGluCYsVal	340
Db	1177	CGTGGAGTCTCATCTGCTGGGAAACAACTGCTGCAAGTCTCTTGTGGAGATGTGTG	1236
OY	341	AspLeuLeuValGlyProGlyValAlaGlnAlaAlaIleLysGluAspProTYrAsnCYs	360
Db	1237	GACCTCTTGtGGGGCCGGGGGCTGCCAGGACGCAATTAAAGAAAGACCTCTGAATCTGC	1296
OY	361	TYrMeTcYsGlyHISlySgLYThrTYrGlyLeuLeuAArgArGluAspTYrProSer	380
Db	1297	TACATGTGCGGGCACAAGGATACCTACCGGGCTGCGCGCGCAGAGAGATCTGGCCCTCC	1356
OY	381	ArgLeuGlnMeTcPhePheAlaAsnAsnHISAspGlnGluPheAspProTolysValTYr	400
Db	1357	CGGCTCCAGATGTTCTTGCTAAATPACACACACACAGCAATTTGACCTCCCAAAGTTTAC	1416
OY	401	ProProValProAlaGluLYsArgLYsProIleArgValLeuSeTYrLeuAspGlyIle	420
Db	1417	CCACTGCTCCAGCTGAGABAAGAGAGGCCATCCGGGGTCTGTCTCTTTGATGGAAATC	1476
OY	421	AlaThrGlyLeuLeuValLeuLYsAspLeuGlyIleGlnValAspArgTYrIIeAlaSer	440
Db	1477	GCTACAGGGCTCTGTGCTGTGAAGACTGTGGCAATCGGGTACGGGGACCGCTACATGGCTCG	1536
OY	441	GluValCYsGluAspSerIIeThrValGlyMeTcValArgHISGlnGlyLYsIIeMeTYr	460
Db	1537	GAGGTGTGTAGACTCATCTACGGTGGGCACTGTCGGCGACACAGGGGAATCATGTATG	1596
OY	461	ValGlyAspValArgSerValThrGlnLYsHISIIeGlnGluTYrGlyProPheAspLeu	480
Db	1597	GTCGGGGAGCTCCGACGCTCACACAGAAAGATATCCAGGATGGGGCCATTGCAATCTG	1656
OY	481	ValIIeGlyLYsTYrProCYsAsnAspLeuSerIIeValAsnProAlaArgLYsGlyLeu	500
Db	1657	GTGATGGGGGCGAGTCCCTGCAATGACTCTTCATCTGTCMAACCTGCTCGAAGGGCTC	1716
OY	501	TYrGlnGlyThrGlyArgLeuPhePheGlnPheTYrArgLeuLeuHISAspAlaArgPro	520
Db	1717	TACGAGGGGCACTGGCGGGCTCTTTGAGTTCTACCGCTCTCTGTCATGATGGCGGCC	1776
OY	521	LYsGlnGlyAspAspArgProPhePheTYrLeuPheGluAsnValValAlaMeGlyVal	540
Db	1777	AAGGAGGGAGATCGCCCTCTTCTTGCGCTTTTGAGAAATGTGTGTGGCATGGCGCTT	1836
OY	541	SeTYrAspLYsArgAspIIeSerArgPheLeuGlnSeTYrAsnProValMeTcIIeAspAlaLYs	560
Db	1837	ACTGCAAGAGGAGCATCTCGGCAATTTCTCGAGTCCAAACCTGTATGATTTGATGCCAA	1896
OY	561	GluValSerAlaAlaHISArgIIaArgTYrPheTYrLYsAsnLeuProGlyMeTcAsnArg	580
Db	1897	GAAAGTGTAGCTGCACACAGGCGCCGCTACTTCTGGGGTAACTCTTCCGGTATGAACGG	1956
OY	581	ProLeuAlaSerThValAsnAspLYsLeuGlnGluGlnGlyCYsLeuGlnHISGlyTYrG	600
Db	1957	CGGTTGGCATTCACCTGTGATATATAGCTGGAGCTGCAGAGATGCTTGAGAGATGGCAG	2016
OY	601	IIeAlaLYsPheSerLYsValArgThrIIeThrTYrArgSerAsnSerIIeLYsGlnGly	620
Db	2017	ATAGCCCAAGTTCAAGCAAACTGAGGACCAATTACTCAAGGTCMAACTCCATTAAGCAGGGC	2076
OY	621	LYsAspArgIIaPheProValPheMeTcAsnGlyLYsGluAspIIeLeuTYrCYsThrGlu	640
Db	2077	AAAGACCAAGCATTTCTGTTCTTCATGAATGAGAAAGAGAACTTTATGTGCTGATA	2136

QY	641	MeculuarYalPheciYpneProvalHnIstYThraspAlaserAmeSerAxlDen	660
Db	2137	ATGGAAGGGATTTGGTTTCCAGTCACCTATACGCTCTCAACATGAGCCGCTTG	2196
QY	661	AlaArgInrArgLeuLeuGlYArgSerTrpSerValProValIleArgHisLeuPheAla	680
Db	2197	GCGAGGCAGACACTGCTGGGCCGTATGAGCCTGCACATGCATCCGACCTCTTGCT	2256
QY	681	ProLeuYsgIurYpheaAlaCysVal	689
Db	2257	CCGCTGAAGGAGTATTTGCGTGTG	2283
RESULT 10			
AF331856	4258 bp	mRNA	linear
LOCUS			PRI 02-JAN-2002
DEFINITION	Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,		
ACCESSION	AF331856		
VERSION	AF331856.1	GI:18033252	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
JOURNAL	Homidae; Homo.		
TITLE	1 (bases 1 to 4258)		
REFERENCE	Ni,J., Pradhan,S. and Roberts,R.J.		
AUTHORS	Cloning, expression and characterization of human DNMT3 genes		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 4258).		
AUTHORS	Ni,J., Pradhan,S. and Roberts,R.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-DEC-2000) New England Biolabs, 32 Tozer Road,		
REFERENCE	Beverly, MA 01915, USA		
FEATURES			
source	Location/Qualifiers		
	1..4258		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="2"		
	1..4258		
	/gene="DNMT3A"		
	227..2936		
	/gene="DNMT3A"		
	/codon_start=1		
	/product="DNA cytosine methyltransferase 3 alpha"		
	/protein_id="AA157039.1"		
	/db_xref="GI:18033253"		
	/translation="MPSGGEIDTSSSAAREDEDKQGEQEPPEPGKEKREPPSTTAR		
	VGRGGRKRKHPVESGDTPKPAVISKSPMAQUSGABELIPNGLEKRSRPOPEEG		
	PAGCGKGAPAEKGAAETLPEAGRAVNGCCTPEKGAGAEAGKEQETNISMKRA		
	AGSNVIRLGGGESSILRORPMRLTPOAGDPYIISKRKDEMLARWREAKAKAV		
	AGNAVDEGRGTELEWTKGRGSSWMPGRIVSMVMGRSAAAGTRVMFGGKFF		
	DEPEYDEGRGTELEWTKGRGSSWMPGRIVSMVMGRSAAAGTRVMFGGKFF		
	VVCEVKMLPSSFCSPFHQATYNTKQPMTRKAIYELVOLYSSRAKLFPVCHDS		
	AKAERKSTAEKPKYKEIIDERTRRLYEVAKOKRNIEDICISGSLNVLLEHLP		
	GMCKNCNCKFLKAYOYDDYOSYCTCCGGRVLMGNNCCRCFCECVLDLVG		
	CAAOAIKEDPMWNCMGHGTGLLRRREMPRLQWFPANNHDOEDFPKVPVPV		
	AEKRPKILVSLFPGIATGLVLVDLGIQVNDYIYASVCEDSITYGAWRHGKIMY		
	DVRSVTOQHIGDEWGPFDLVITGSPNDISIVNPKRGLYESTGRLFPEFYRLH		
	KEGDRPEFMLEFENVVAVGSDKRDISFLESNSVMIDAKVSAHRAKYFMGNIP		
	NRPLASTVNDKLEOECLENHRIAKFSKVRTITTRSNSIKGKQOHFVFNNEK		
	WCTEMERVFGEFHYTIVDSNMSRLARQLRLGRSWSVPVIRHLFAPLKEYFACV"		
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	4258
Score:	3714.00	Matches:	679
Percent Similarity:	98.8%	Conservative:	2
Best Local Similarity:	98.5%	Mismatches:	8
Query Match:	98.8%	Indels:	0
DB:	5	Gaps:	0

US-10-623-813-85 (1-689) x AF331856 (1-4258)

Oy 1 MetAsnAlaValaGluGluAsnGlnAlaSerGlyGluSerGlnLysValaGluGlnAlaSer 20
 Db 887 ATGAATGCTGTGGAGAAAAACCGGGGCCCGGGGAGCTCTAGAAAGCTGAGGAGCGCCAC 946
 Oy 21 ProProAlaValaGluGlnProThrAspProAlaSerProThrValaAlaThrThrProGlu 40
 Db 947 CTTCTGCTGTGTGACAGCCCACTGACCCCGCATCCCACTGTGCTACACAGCTTGAG 1006
 Oy 41 ProValaGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 Db 1007 CCCGTGGGGGTCCGATGCTGTGGGACAAAGATGCCCAAGCAGGCGATACAGAGCCAGAG 1066
 Oy 61 TyrGluAspGlyAArgGlyPheGlyTlleglyGluLeuValaTrrGlyLysLeuArgGlyPhe 80
 Db 1067 TACGAGGACCGCGCGGGCTTTGGCATTTGGGAGCTGTGTGGGGGAAACTGCGGGGCTTC 1126
 Oy 81 SerTrrProGluAArgLysValaSerTrrProMetThrGlyAArgSerArgAlaAlaGlu 100
 Db 1127 TCTGTGTGCGCAGCGCGCATTTGTTGTGTGATGACGGCGCGAGCCGACAGACTGAA 1186
 Oy 101 GlyThrArgTrrValaMetTrrPheGlyAAspGlyLysPheSerValaValaCysValaGluLys 120
 Db 1187 GGCACCCCGCTGGGTATGTGTGTCGAGAGACGGCAAACTCTCAGTGTGTGTGTGAGAG 1246
 Oy 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 Db 1247 CTGATGCGCGTACGCTCGTTTGTGAGTGGCTTCCACAGGCGACGTAACAAGCAGCC 1306
 Oy 141 MetTrrArgLysAlaTlleglyGluValaLeuGlnValaAlaSerSerArgAlaGlyLysLeu 160
 Db 1307 ATGTAACCCCAAGCCATCTACAGAGTCTTCGAGGTGCGACAGCCGCGCGGAGAGCTG 1366
 Oy 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaAlaGluValaGlnAsn 180
 Db 1367 TTCCCGGTGTGCGACAGACGCACTGAGTGCACCTGCCAAGCCCTGGAGGTGCAGAAC 1426
 Oy 181 LysGlnMetTlleglyTrrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGln 200
 Db 1427 AAGCCCATGATTGATGGGCGCTTGGGGGGCTTCCAGCCTTTCGGCCCTAAGGGCCCTGAG 1486
 Oy 201 ProProGluGluGlnLysAsnProTrrLysGluValaTrrThrAspMetTrrValaGluPro 220
 Db 1487 CCACAGAGAGAGAGAGATCCCTACAAAGAGTACACGAGCATGTGGGTGAACT 1546
 Oy 221 GluAlaAlaAlaTrrAlaProProProProAlaLysLysProArgLysSerThrTrgln 240
 Db 1547 GAGGAGCTGCTTACGACCACTTCCACAGCCCAAAAGCCCCGAGAGGACACAGCGAG 1606
 Oy 241 LysProLysValaLysGluTlleglyLeaAspGluArgTrrArgGluValaTrrGluVala 260
 Db 1607 AAGCCCAAGGTCAAGAGATTATTGATGAGCCACAAAGAGACGGCTGTGTACGAGGTG 1666
 Oy 261 ArgGlnLysCysArgAsnTlleglyLysPrlleCysTlleserCysGlySerLeuAsnValaThr 280
 Db 1667 CGGCGAAGTGCAGAAATTAGAGACATCTCATCTCTGTGGAGCTCTCATGTATACC 1726
 Oy 281 LeuGlnHisProLeuPheTlleglyLysMetCysGlnAsnCysLysAsnCysPheLeuGln 300
 Db 1727 CTGGAACACCCCTCTTCTGTGTGAGAGAAATGTCCCAAACTGCAGAAACTCTTCTGAG 1786
 Oy 301 CysAlaTrrGlnTrrAspAspAspGlyTrrGlnSerTrrCysThrTllegCysCysGlyGly 320
 Db 1787 TGTGGGTACAGTACGACGACGCGCTACAGTCTTACTGACACATCTGTGTGGGAGC 1846
 Oy 321 ArgGluValaLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValaGluCysVala 340
 Db 1847 CGTGAGGTGCTCATGTGCGGAACAACATCTGTGAGGTGCTTTTGTGCTGAGAGTGTG 1906
 Oy 341 AspLeuLeuValaGlyProGlyAlaAlaGlnAlaAlaTlleglyGluAspProTrrAsnCys 360

Db 1907 GACCTTTGTGTGGGCGCGGGGCTGCCAGGACGACCATTAAGAAAGACCTTCGAACCTGC 1966
 Oy 361 TyrMetCysGlyHisLysGlyTrrTrrGlyLeuLeuAArgAArgGluAspTrrProSer 380
 Db 1967 TACATGTGCGGGACAAAGGTATCTTACGAGCTGTGCGCGCGCGAGAGGACTTGGCCCTTC 2026
 Oy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValaTrr 400
 Db 2027 CGGCTCCAGATTTCTTCTGCTATTAACACAGACAGAAATTTGACCTTCCAAAGCTTAC 2086
 Oy 401 ProProValaProAlaGluLysArgLysProLysArgValaLeuSerLeuPheAspGlyTl 420
 Db 2087 CCACCTGTCCAGCTGAGAAAGAGAACCCATCCGGGTGTCTCTCTTATGGAATC 2146
 Oy 421 AlaThrGlyLeuLeuValaLeuLysAspLeuGlyTlleglnValaAspArgTrrTllegAsn 440
 Db 2147 GCTACAGGAGCTCTGTGTCTGAAAGCATTTGGCATTCAGGTGACCGGCTACATTCCTG 2206
 Oy 441 GluValaCysGluAspSerTllegThrValaGlyMetValaArgHisGlnGlyLysIleMetTyr 460
 Db 2207 GAGTGTGTGAGACTTCATCACGCTGGCATGTGTCGGCACCAAGGGAAGATCATGTAC 2266
 Oy 461 ValGlyAspValaArgSerValaThrGlnLysHisTlleglnGluTrrGlyProPheAspLeu 480
 Db 2267 GTGGGAGCGTCCGACGCTCACAGAAAGCATATCCAGAGTGGGCCCATTCGATCTG 2326
 Oy 481 ValTlleglyLysSerProCysAsnAspLeuSerTllegValaAsnProAlaArgLysGlyLeu 500
 Db 2327 GTGATTTGGGGAGTCCCTGCAATGACTTCCATGTCTCAACCTGTGCAAGGGCTTC 2386
 Oy 501 TyrGlnGlyThrGlyAArgLeuPheGluPheTrrArgLysLeuHisAspAlaArgPro 520
 Db 2387 TACGAGGCACTGGCGGCTCTCTTGAATGTTCTACCCCTCTCATATGTCGGGCCC 2446
 Oy 521 LysGlnGlyAspAspArgProPhePheTrrLeuPheGluAsnValaAlaMetGlyVala 540
 Db 2447 AAGAGGAGATGATCGCCCTTCTTGTGCTTGTGAATGTGTGGCCCATTCGGCGCTT 2506
 Oy 541 SerAspLysAlaArgAspTlleserArgPheLeuGluSerAsnProValaMetLysAspAlaLys 560
 Db 2507 AGTGCAAGAGGACATCTCGGATTTCTCGATCCCAACCTGTATGATTATGCTCAA 2566
 Oy 561 GluValaSerAlaAlaHisArgAlaArgTrrPheTrrGlyAsnLeuProGlyMetAsnArg 580
 Db 2567 GAGTGTGAGCTGCACAGAGGCGCGCTACTTCTGGGTACCTTCCCGTATGAACAG 2626
 Oy 581 ProLeuAlaSerThrValaAsnAspLysLeuGlnGlnCysLeuGlnHisGlyArg 600
 Db 2627 CCGTTGGCATCCACTGTAATGATTAAGCTGAGCTGCAGAGTGTCTGGACATGGCAGG 2686
 Oy 601 TllegAlaLysPheSerLysValaArgTrrTllegThrArgSerAsnSerTllegGlnGly 620
 Db 2687 ATAGCAAGTTCAGAAAGTGAAGCACTTACTAGAGGTCAAACTCCATAAAGACGGGC 2746
 Oy 621 LysAspGlnHisPheProValaPheMetAsnGluLysGluAspTllegTrrCysThrGln 640
 Db 2747 AAGACCGCATTTTCTCTCTCTTCAATGATGAGAAAGAGACATTTATGTGTGACTGA 2806
 Oy 641 MetGluArgValaPheGlyPheProValaHisTrrThrAspValaSerAsnMetSerArgLeu 660
 Db 2807 ATGAAAGGATATTGTGTTTCCAGTCCACTATATGACGTCTCCAAACATGAGCGGCTTG 2866
 Oy 661 AlaArgGlnArgLeuLeuGlnGlyAArgSerTrrSerValaProValaTllegArgHisLeuPheAla 680
 Db 2867 GCGAGGAGAGACGTGTGGCGGTATGAGAGCGGTGCCAGTATCCGACCTCTTCGCT 2926
 Oy 681 ProLeuLysGluTrrPheAlaCysVala 689
 Db 2927 CCGTGAAGAGATATTGCGGTGTG 2953

RESULT 11
 DD171352
 LOCUS

DD171352 4294 bp DNA linear PAT 19-DEC-2005

DEFINITION Marker for diagnosing human subarachnoid hemorrhage and its use.
 ACCESSION DDI71352
 VERSION DDI71352.1 GI:83959786
 KEYWORDS JP 2005151854-A/10.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 4294)
 Yamaura, A., Kubota, M., Machida, T., Uchino, F., Kobayashi, E., Sasaki, N., Takiguchi, M. and Hiwasa, T.
 Marker for diagnosing human subarachnoid hemorrhage and its use
 Patent: JP 2005151854-A 10 16-JUN-2005;
 Japan Science and Technology Agency
 COMMENT OS Homo sapiens
 PN JP 2005151854-A/10
 PD 16-JUN-2005
 PF 21-NOV-2003 JP 2003393161
 PI akira yamaura,mocoo kubota, toshio machida, fukuo uchino, eichi
 PI kobayashi,
 PI nakatsu saeki, masaki takiguchi, takaki hiwasa CC
 FH Key Location/Qualifiers.
 FEATURES
 source 1..4294
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 4294
 Score: 3714.00 Matches: 679
 Percent Similarity: 98.8% Conservative: 2
 Best Local Similarity: 98.5% Mismatches: 8
 Query Match: 98.8% Indels: 0
 Gaps: 0
 DB: 2
 US-10-623-813-85 (1-689) x DDI71352 (1-4294)
 QY 1 MetAsnAlVAAlGluGluAEngInAlAsErGlyGluSerGlnYsValGluGluAlAsEr 20
 DB 907 ATGAATGCTGTGTAAGAAACCAAGGGGCCCGGGAAGTCTCAGAAGGTGAGGAGCCAGC 966
 QY 21 ProProAlaValGlnGlnProThraSPProAlaSerProThraValAlaThrTrpProGlu 40
 DB 967 CCGCTCGCTGTGACGAGCCCACTGACCCCGCATCCCGCATCTGTGTACACGCGCTGAG 1026
 QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 1027 CCGGTGGGTCCGATGCTGGGACAGAAATGCCAACCAAGCAGCGATGACGAGCCAGG 1086
 QY 61 TYrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
 DB 1087 TACGAGAGCGCGCGGGCTTTGGCATTGGGAGCTGTGTGGGAAACTGCGGGGCTTC 1146
 QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 DB 1147 TCCTGTGGCCAGAGCCGATTTGTCTTGTGTGATGACGCGCGGAGCCGAGCAGCTGAA 1206
 QY 101 GlyThrArgTrpAlaMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 1207 GGCACCCGCGGGATCATGTGTTGGAGAGAGCAAAATTCATGAGGTGTGTGAGAG 1266
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyraSnlYsgInPro 140
 DB 1267 CTGATCCCGCTGAGCTCGTTTTCAGATGCTTCCACAGGCAAGCTATCAACAGAGCC 1326
 QY 141 MetTyraArgLysAlaIleTyrgluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 1327 ATGTACCGCAAAACCATCTACGAGGTCTCTCAGGTGCGCAGCGCCCGGGGAAAGCTG 1386
 QY 161 PheProAlaCysHisAspSerAspGluSerCysPserGlyLysAlaValGluValGlnAsn 180

DB 1387 TTCCCGGTGCGCACGACGATGAGTGAACCTGCCAAGGCGGTGAGGTGCGAAGC 1446
 QY 181 LysGlnMetCileGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 DB 1447 AAGCCCATGATGAATGAGGCGCTTCCAGCCCTTCTGCGCCCTTAAGGCGCTTGAG 1506
 QY 201 ProProGluGluGluLysAsnProTyrgluValIleTyThrAspMetTrpValGluPro 220
 DB 1507 CCACCAAGAAAGAGAAATCCCTACAAAGAAAGTACACGACATGTGGTGAACCT 1566
 QY 221 GluAlaAlaAlaTyraLapProProProAlaLysLysProArgLysSerThrGlu 240
 DB 1567 GAGGAGCGTGGCTTACGCAACACTCCACGACCAAAAGAGAGAGAGAGAGAGAGAG 1626
 QY 241 LysProLysValLysGluIleLeaSPGluAgtThrArgGluArgLeuValIleTyrgluVal 260
 DB 1627 AAGCCCAAGTCAAGAGATTAATGATGAGCCACAAAGAGAGAGCGCTGTGTACGAGGTG 1686
 QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1687 CGCAGAAAGTCCGGAAACATTAAGACATCTGCATCTCTGTGGAGCTCAATGTACC 1746
 QY 281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 DB 1747 CTGAAACACCCCTCTCTGCTGTGAGGAATGTGCAGAAATCGCAAGAACTGCTTCTGAG 1806
 QY 301 CysAlaTyrgluTyraSPAspGlyTyrglnSerTyrglnSerThrIleCysGlyGly 320
 DB 1807 TGTGCTTACAGTACACACACAGCGCTACAGTCTTACTGACCACTGTGTGTGGGGG 1866
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 DB 1867 CGTAGAGTCTCATGTGCGGAAACAAACACTGCTCAGGTGCTTTGCTGAGAGTGTG 1926
 QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
 DB 1927 GACCTTTGTGTGGGGCGGGGCTGCCAGGACGACATTAAAGAAAGCCCTGGAACCTGC 1986
 QY 361 TyrMetCysGlyHisLysGlyThrTyrglyLeuLeuArgArgGluAspTrpProSer 380
 DB 1987 TACATGTGGGACACAAAGGTACTTACGGGCTGTGGGCGGAGAGAGAGAGAGAGAGAG 2046
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValIle 400
 DB 2047 CGGCTCAATGTTCTTCTCTTAATTAACACGACAGAAATTTGACCTCCAAAGTTTAC 2106
 QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2107 CCACTGTCCAGCTGAGAAAGAAAGCCATCCGGGTGCTGTCTCTTGTGATGAATC 2166
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrlleAlaSer 440
 DB 2167 GCTACAGGCTCTCTGTGTGAAGACTTGGGCACTTCAAGTGAACGCTACATTTGCTTCG 2226
 QY 441 GluValCysGlnAspSerIleThrValGlyMetValArgHisGlnGluTrpGlyProPheAspLeu 460
 DB 2227 GAGGTGTGAAGACTTCATCACGTTGGGCAATGTGCGGCAACAGGGAAGATCATCTAC 2286
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
 DB 2287 GTCCGGGAGCTCCGAGCTGCACACAGAAAGCATATCAAGAGGAGGCGCCATTCGATCG 2346
 QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2347 GTGATTTGGGGAGCTCCCGCAATGACCTCTCATCTGTAACCTCTGCGCAAGGCGCTC 2406
 QY 501 TyrGluGlyThrGlyArgLeuPhePheGlyPheTyArgLeuLeuHisAspAlaArgPro 520
 DB 2407 TACGAGGAGACTGGCGGCTCTTCTTGAAGTTCTACCGGCTCTGTGATGATGCGGGCC 2466
 QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540

Db 2467 AAGGAGGAGATGATGCGCCCTTCTTCTGCTCTTTGAGATGCTGGCCATGAGGCTT 2526

Qy 541 SerAspLysArgAsp111SerArgPheLeuGluSerAsnProValMet111LeaSPAlaLys 560

Db 2527 AGTGACAGAGGAGACATCTCGCATTTCTCGAGTCCAAACCTGTGATGATGATGCCAA 2586

Qy 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGluAsnLeuProGlyMetAsnArg 580

Db 2587 GAAGTGTCAAGCTCACACAGGCGCCGTACTCTGGGGTAACCTTCCCGGTATGACAGG 2646

Qy 581 ProLeuAlaSerThrValaAsnAspLysLeuGluLeuGlnGluCysLeuGlnIun1sglyArg 600

Db 2647 CGTTGGCATCCACTGTGAATGATTAAGCTGAGCTGAGAGAGTGTCTGAGCATGGCAGG 2706

Qy 601 11eAlaLysPheSerLysValaArgThr11eThrThArgSerAsnSer11elyGlnGly 620

Db 2707 ATAGCCAAAGTTCAGCAAGAGGAGACCATTAACGAGTCAAACTCCATAAAGCAGGCGC 2766

Qy 621 LysAspGlnHisPheProValaPheMetAsnGluLysGluAsp11eLeuTrpCysThrGlu 640

Db 2767 AAGAGCAGCAGATTTCTCTGCTTCTCATGATGAGAAAGAGACATCTTATGTGCACTGAA 2826

Qy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660

Db 2827 ATGGAAAGGATATTGGTTTCCAGTCCATATAGAGCTGCCAACAATGAGCCGCTTG 2886

Qy 661 AlaArgGlnArgLeuLeuGlyArgSerTyrPseValProVal11eArgHisLeuPheAla 680

Db 2887 GCGAGCAGACAGACGCTGGCGCCGCTCATGAGAGCGTGCAGTCCGACCTCTTCGCT 2946

Qy 681 ProLeuLysGluTyrPheAlaCysVal 689

Db 2947 CCGCTGAGAGATATTGCGTGTGTG 2973

RESULT 12
BC043617 4294 bp mRNA linear PRI 30-JUN-2004
LOCUS BC043617
DEFINITION Homo sapiens DNA (cytosine-5'-methyltransferase 3 alpha,
transcript variant 3, mRNA (cDNA clone MGC:50948 IMAGE:6150112),
complete cds.
ACCESSION BC043617 GI:27694443
VERSION BC043617.1
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 4294)
Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
Dietzenko L., Martusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L.,
Schaefer T.E., Brownstein M.J., Usdin T.B., Teshiyuki S.,
Carinci P., Prange C., Kana S.S., Loquellano N.A., Peters G.J.,
Adams R.D., Mullany S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S.,
Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E.,
Scherer A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL
PUBMED 12477932
REFERENCE 2 (bases 1 to 4294)
AUTHORS Straussberg R.

TITLE
JOURNAL
DIRECT SUBMISSION
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCID/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Buterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liso,
Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu,
Paraneh Saeedi, JR Santos, Angelique Scherch, Ursula Skalska,
Duane Smalins, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.nih.gov>
Series: IRAC Plate: 89 Row: 0 Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency OR
analysis.
FEATURES
source
Location/Qualifiers
1..4294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:50948 IMAGE:6150112"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH MGC_72"
/lab_host="DH10B"
/note="vector: PCMV-SPORT6"
1..4294
/gene="DNMT3A"
/note="synonyms: DNMT3A2, M.Hsa111A"
/db_xref="GeneID:1788"
/db_xref="MIM:602769"
238..2976
/gene="DNMT3A"
/codon_start=1
/product="DNA cytosine methyltransferase 3 alpha, isoform
a"
/protein_id="AAH43617.1"
/db_xref="GI:27694444"
/db_xref="GeneID:1788"
/db_xref="MIM:602769"
/translation="MPAMPSSGCPDPTSSAAREEDRKDGEORERKGEORPESTT
AKVGRPKRGKRPVSGDTPKPAVKSISMSQDSSASELIPNGDEKSEPOE
EGSPAGGKGAPAGGGAETLPPASRAVEGSCCKPKRGAPAGGAEKENTIES
KMGESRGRLRGKLGESLIRORPMPRLTPOGDPVYSKRRDEMLARKGEAEKA
KTIAGNNAAVEENGQGESEOKVEASPPAQOTPDASPTVATTPPEVGSADKDKATK
AADPEPEDEGFGIGELVWKLKLFSSWMPRIIVSWMTGSRSAEGTRVWVWFEDG
KRSVAVCEKMLPSFSCSAFHQATYKQKPYKAIYEVQVASSRAGKLPVCHDSE
SSTAAVAYONKEMIALGFGPSPGPKLSEPKRNPKYKXYTDMWPEPAAYAP
PPPAKKPRKSTAEKRVKELIDERTREIRLVYEVROKRNKIEDICISGLVNTLEPL
PFGKCKNCKNCKNCKNCAVOYDDGYOSYCTTICGGREVMCGNNNCGRCFVCEVDL
VPGAAQAALIKEDPNCYCKGKTYGILARRDEDPSTROMEPANNDPEDPKRYP
EPVPAKRPKRIYLSFDGATLVLVKDGLGVDRYASVECDSTTVGMRRQKYM
VYGDVRSYQKIQSGEPDVLIGSPCNDSLIVNPARGLYEGTGLFFETRLHD
AAPKGDPRPFPMLENNVAVMGVSDKRDLSRPLESPNVMIDAKEVSAARARFVWNL
PQANRPPLASTVNDKLEOCELEHGRILAKSKRTITRTRENSIKQGDQHPVFMNEKE
DILKCTEMERVGPVHYTDVSNMRLARQRLGSSWSVPVIRHLPAFLKEYACV"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4294

Score:	3714.00	Matches:	679
Percent Similarity:	98.8%	Conservative:	2
Best Local Similarity:	98.5%	Mismatches:	8
Query Match:	98.8%	Indels:	0
DB:	5	Gaps:	0
US-10-623-813-85 (1-689) x BC043617 (1-4294)			
QY	1	MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnLysValGluGluAlaSer	20
DB	907	ATGAAATCTGTGGAAAGAAAACCAAGGGCCCGGAGAGTCTCAGAAAGTGGAGAGCCAGC	966
QY	21	ProProAlaValGlnGlnInProThraSProlaSerProThraValAlaThrThraProGlu	40
DB	967	CTCTCTCTGTGACGACGACCCACTGACCCGATCCCTCCACTGGCTGCTACCACTGAG	1026
QY	41	ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu	60
DB	1027	CCGGTGGGTCCGATGCTGGGGACAGAAATCCCAAGAGCGCATGACGAGCCAGAG	1086
QY	61	TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValIleGlyLysLeuArgGlyPhe	80
DB	1087	TACGAGACGCGCGGCTTTGGCATTTGGGAGCTGGTGTGGGGAATCTGGCGGCTTC	1146
QY	81	SerTrpTrpProGlyArgGlyIleValSerTrpTrpMetThrGlyArgSerArgAlaGlu	100
DB	1147	TCTGTGGGCGAGGCGCATTTGTCTTGGTGAAGAGGCGGAGCCGAGCCGAGCTGAA	1206
QY	101	GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys	120
DB	1207	GGCACCGCGTGGGTCACTGGTTGGAGAGCGGCAAAATTCACAGTGTGTGTGAGAAAG	1266
QY	121	LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTrpAsnLysGlnPro	140
DB	1267	CTGATGCGCGTGGCTGTGTGTTCAGTGTCTCCACAGGCGCACTGCAACAGCACGCC	1326
QY	141	MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu	160
DB	1327	ATGTACCGCAAGACCATCTACAGAGTCTCTCAGGTGCGCAGACCGCGCGGGAGCTG	1386
QY	161	PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn	180
DB	1387	TTCGCGGTGTGCCACGACGAGTGAAGTACACTGCCAAGGCGGTGAGTGCAGAAC	1446
QY	181	LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu	200
DB	1447	AACCCCATGATTGAAATGGGCTGTGGGGCTTCCAGCTTCTGGCCCTTAAGGGCTGGAG	1506
QY	201	ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro	220
DB	1507	CCACCGAAGAGAGAAATCCCTACAAAGATGTAACAGGACATGTGGTGGAACTT	1566
QY	221	GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu	240
DB	1567	GAGGAGCTGCTCACCACCACTCCACAGCCAAAGCCCGGAGAGACACAGCGGAG	1626
QY	241	LysProLysValLysGluIleLysPheGluArgThrArgGluArgLeuValTyrGluVal	260
DB	1627	AACCCCAAGGTCAAGAGATTATTGATGACGCAAGAGAGCGGCTGGTGTGAGAGTG	1686
QY	261	ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr	280
DB	1687	CGGAGAGAGTCCGAGCATTTGAGACATCTGCTCTGTGGAGCTCCAAATTTACC	1746
QY	281	LeuGlnHisProLeuPheIleGlyLysMetCysGlnAsnLysLysAsnLysPheLeuGlu	300
DB	1747	CTGGAACACCCCTCTTGTGGAGAAATGTGCCAAAGCTGCAAGAACTCTTCTGAG	1806
QY	301	CysAlaIleTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly	320
DB	1807	TGTGCGTACAGTACGACGAGAGGCTTACATGCTTCACTGCAACATCTCTGTGGGGG	1866
QY	321	ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal	340

DB	1867	CTGAGGTCTCATGTGCGGAAACAACTGTGTGAGAGTCTTTTGGTGGAGTGTG	1926
QY	341	AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProThraPheCys	360
DB	1927	GACCTCTTGTGGGGCCGGGGGCTGCCAGGAGCCATTAAAGAAACCCCTGGAATCG	1986
QY	361	TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer	380
DB	1987	TACATGTGGGACAAAGGTACTTACGAGCTGTGGCGGCGGAGAGAGCTGGCCCTCC	2046
QY	381	ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr	400
DB	2047	CGGCTCAGATGTTCTTCCCTTAATTAACAGACAGCAAGATTTGACCTCCAAAGTTTAC	2106
QY	401	ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle	420
DB	2107	CCACTGTCCCACTGAGAAAGAGAAAGCCCATCCGGGTCTGTCTTGTGATGGAATC	2166
QY	421	AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer	440
DB	2167	GCTACAGGCTCTGTGTGCTGAAGACTTGGGCATTCAAGTGAACGCTACATGGCTCG	2226
QY	441	GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr	460
DB	2227	GAGGTGTGAGAGACTCCATCACGCTGGGCATGTGCGCACAGGGGAAAGATCATGTAC	2286
QY	461	ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu	480
DB	2287	GTCGGGAGAGCTCCGCGCTCACACAGAAATATCCAGAGAGGGGCCCATTCGATCTG	2346
QY	481	ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu	500
DB	2347	GTAATTGGGAGAGTCTCTGCAATGACTTCTCATGTCAACCTCTCTGCAAGGCTTC	2406
QY	501	TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro	520
DB	2407	TACGAGGACATGGCGGCTCTTCTTGAATTCTACCGCTCTGATGATGGCGGCC	2466
QY	521	LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal	540
DB	2467	AAGGAGGAGAGATGCCCCCTTCTTGTGGCTTGTGAATGTGTGGCCCAAGGGCGTT	2526
QY	541	SerAspLysArgAspIleSerArgPheLeuGlnLysSerAsnProValMetIleAspAlaLys	560
DB	2527	AGTGAAGAGGAGACTCTCGGATTTCTCGAGTCCAACTGTGATGATGATGCCAA	2586
QY	561	GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg	580
DB	2587	GAAGTCTCAGTGCACACAGGGCCGCTACTTCTGGGGTAACTTCCCGTATGAACAG	2646
QY	581	ProLeuAlaSerThrValAsnAspLysLeuGlnGluCysLeuGlnHisGlyArg	600
DB	2647	CCGTTGGCATTCACCTGTGAATATTAAGCTGACAGCTGCAGAGAGTCTTGAGACATGG	2706
QY	601	IleAlaLysPheSerLysValArgThrIleThrArgSerAsnSerIleLysGlnGly	620
DB	2707	ATAGCCAAAGTTCAGCAAGATGAGACATTACTACAGGTCAAATCTCCATTAAGCAGGG	2766
QY	621	LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu	640
DB	2767	AAAGACACACATTTCTCTCTTCAATGAATGAAGAAAGACATCTTATGTGTACATGAA	2826
QY	641	MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu	660
DB	2827	ATGGAAGGTATTTGGTTCCAGTCCACTATACGATGCTCCCAACATGAGAGCCCTTG	2886
QY	661	AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla	680
DB	2887	GCGAGGACAGACTGTGGGCGGCTCATGAGCGTGCAGTCAATCCGACACCTCTTGCT	2946
QY	681	ProLeuLysGluTyrPheAlaCysVal	689

Db 2947 CCAGTGAAGAGTATTTGGCTGTG 2973

RESULT 13
CS050241 4395 bp DNA linear PAT 23-MAR-2005
LOCUS CS050241
DEFINITION Sequence 25 from Patent WO2005021757.
ACCESSION CS050241
VERSION CS050241
KEYWORDS GI:61889496
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1
AUTHORS Tomme, P. H. and van Rompaey, L.
TITLES Polypeptides and polynucleotides for use as a medicament
JOURNAL Patent: WO 2005021757-A 25 10-MAR-2005;
Galapagos Genomics N.V. (BE)

FEATURES
source Location/Qualifiers
1..4395
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 4395
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
DB: 2 Gaps: 0

US-10-623-813-85 (1-689) x CS050241 (1-4395)

QY 1 MetAsnAlaValAGluGluAsnGlnAlaSerGlyGluSerGlnValAGluGlnAlaSer 20
Db 1008 ATGAAATGCTGTGGAAACCAAGGGGCGGGGAGTCTCAAGAGTGAAGAGGGCCAGC 1067
QY 21 ProProAlaValAGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db 1068 CCTCGGCTGTGAGAGCCAGCCAGTCCCGCATCTCCCACTGTGCTACAGCTGAG 1127
QY 41 ProValAGlyValAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
Db 1128 CCGGTGGGCTCCGATCTCTGGGGGACAAAGATGCCACCAAGCAGCGCATACAGCGCCAGAG 1187
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyLysLysValTyrGlyLysLysLysPhe 80
Db 1188 TAGAGAGAGCGCGGGGCTTTGGCATTTGGAGACTGTGTGGGGAACCTGCCGGGCTTC 1247
QY 81 SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db 1248 TCCTGGTGGCCAGCGCGCATTTGTCTTGTGTGATACCGGGCCGAGCCGAGAGCTGAA 1307
QY 101 GlyThrArgTyrValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
Db 1308 GGCACCCCGCTGGGTCTGTGTGTCGAGACGGCAAAATCTCACTGTGTGTGTGAGAG 1367
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db 1368 CTGATGCCCTGAGCTCGTTTGGCACTGGCTTCACACAGGCCACGTAACAACAGAGCC 1427
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db 1428 ATGTACCCGAAAGCCATCTACAGAGTCTGCAAGTGGCCAGCGCGCGGGAAGCTG 1487
QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAn 180
Db 1488 TTCCCGGTGTGCCACACAGCGATGAGATGACATGCCCAAGCCCTGTGAGGTGCGAAGC 1547
QY 181 LysGlnMetIleGlnTyrPheLysLysGlyLysPheGlnProSerGlyProLysGlyLysGlu 200

Db 1548 AAGCCCATGATTGAATGGCCCTGGGGGGCTTCAGACCTTCTGGCCCTAAGGCGCTGAG 1607
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db 1608 CCACAGAAAGAAAGAGAAATCCCTCAAAAGAGTGTACACGACATGTGGGTGAACCT 1667
QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
Db 1668 GAGGAGCTGCTTACGACACCTTCCACAGCCAAAGAGCCCGGAGAGCAGCGGAG 1727
QY 241 LysProLysValLysGluIleIleAspGluArgTyrArgLysValTyrGluVal 260
Db 1728 AAGCCCAAGGTCAAGAGATTATTGATGAGCCACAAAGAGAGCGGCTGTGTACAGGTG 1787
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 1788 CGCAGAAAGTGGCGGAACATTGAGACATCTGCATCTCTGTGGAGCCTCAATGTTAC 1847
QY 281 LeuGlnHisProLeuPheIleGlyLysMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db 1848 CTGAAACACCCCTCTTCTGTGAGAGATGTCCAAACTGCAAGACTGCTTCTGGAAG 1907
QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1908 TGTGCGTACCACTACAGACAGACAGCGCTACAGTCTCTACACCATCTGTGTGGGGC 1967
QY 321 ArgGluValLeuMetCysGlyLysAsnAsnAsnCysArgArgCysPheCysValGluCysVal 340
Db 1968 CGTGAAGTGTCTATGTGTGGGAAACAACATCTCAGGTGTCTTTGCTGTGAGTGTG 2027
QY 341 AspLeuLeuValAGLProGluValAlaGlnAlaIleLysGluAspProTrpAsnCys 360
Db 2028 GACCTTGTGTGGGCGGGGCTGCCAGGACCATTAAGAAACCCCTGGAACTGC 2087
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
Db 2088 TACATGTGGGCAACAAGGATCTTACGGGCTGTGGCGGCGGAGAGAGATGGCCCTCC 2147
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db 2148 CGGTCTCCAGATTTCTTGTCTAATTAACACAGCACAGAAATTGACCTTCAAGGTTTAC 2207
QY 401 ProProValProAlaGluLysAspLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 2208 CCACCTGTCCAGCTGAGAGAGAGGCCATCCGGGTGCTCTCTTTGATGGAAATC 2267
QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 2268 GCTACAGGGCTCTGTGTGTGAAGACTTGGGCACTTCAAGGTGACCGGCTACATTGGCTCG 2327
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 2328 GAGGTGTGTGAGGATCTTCATCACTGTGGGCACTGTGGCGGACAGGGGAAGTCAATGTATC 2387
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
Db 2388 GTGGGGAGCTCCGACAGGTCAACAGAACTATTCAGAGATGGGGCCATTCCGATCTCG 2447
QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAspProAlaArgLysGlyLeu 500
Db 2448 GTGATGGGGGAGAGCTCTGCAATGACTCTTCATCTGCAACCTGCTGCAAGGGCTTC 2507
QY 501 TyrGluGlyThrGlyArgLeuPheGluPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 2508 TACGAGGGGACATGGCGGCTCTTCTTGAATTCACCGCTCTCTCATATATCCGGGGCC 2567
QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db 2568 AAGGAGGAGATGATGCCCTTCTTGTGGCTTTTGAGAAATGTGTGGCCATTGGGCTT 2627
QY 541 SerAspLysArgAspIleSerArgPheLeuGluLeuSerAsnProValMetIleAspAlaLys 560

Db 2628 AGTGACAGAGGACATCTCGCATTTCTGAGTCCACCTGTGATTGATTCACAA 2687
 QY 561 GUAValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 560
 Db 2688 GAAGTGTCACTGACACACAGGCGCCGCTACTTGTGGGTAACTTCCCGGTATGAACAGG 2747
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnGlnLucysLeuGlnLysArg 600
 Db 2748 CCTTGGCATCCACTGTGAATGATTAAGCTGAGCTGACGAGAGTGTGGAGATGGCAGG 2807
 QY 601 LLeAlaLysPheSerLysValArgThrLLeThrThrArgSerAsnSerLLeLysGlnGly 620
 Db 2808 ATAGCCACAGTTCACAAAGTGAAGACCATTAACAAGTCAAACTCCATAAAGCAGGCGC 2867
 QY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspLLeLeuTyrCysThrGln 640
 Db 2868 AAAGACACAGCATTTCTCTGTCTTCAATGATGAGAAAGAGACATCTTATGTGTCACATGAA 2927
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspAlaSerAsnMetSerArgLeu 660
 Db 2928 ATGAAAGGATATTGGATTTCACAGTCCACTATACGACGTCTCCAAACATGAGCCGCTTG 2987
 QY 661 AlaArgGlnArgLeuGlnArgSerTyrSerValProValLLeArgHisLeuPheAla 680
 Db 2988 GCGAGGCAAGACTGCTGGGCGCGTCAATGAGCGTGCACATCCGCCACCTCTTCGCT 3047
 QY 681 ProLeuLysGlnTyrPheAlaCysVal 689
 Db 3048 CCGCTGAGAGATTTTGGCTGTGTG 3074
 RESULT 14
 CQ720469 2841 bp DNA 1linear PAT 03-FEB-2004
 LOCUS CQ720469 Sequence 6403 from Patent WO02068579.
 DEFINITION CQ720469
 ACCESSION CQ720469
 VERSION CQ720469.1 GI:42281326
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 REFERENCE 1
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 TITLE Kites, such as nucleic acid arrays, comprising a majority of
 humenexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 6403 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES
 source Location/Qualifiers
 1..2841
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2841
 Score: 3697.00 Matches: 679
 Percent Similarity: 96.9% Conservative: 2
 Best Local Similarity: 96.6% Mismatches: 8
 Query Match: 98.4% Indels: 14
 Db: 2 Gaps: 1
 US-10-623-813-85 (1-689) x CQ720469 (1-2841)
 QY 1 MetAsnAlaValAlaGlnGlnGlnAlaSerGlyLysSerGlnLysValGlnGlnAlaSer 20
 Db 730 ATGAATGCTGTGAAAGAAACCAAGGGCCCGGGAGTCTCAGAAAGTGGAGAGGCCACG 789
 QY 21 ProProAlaValAlaGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 Db 790 CCTCTGCTGTGAGGAGGACCATGACCCCGCATCCGCCACTGTGGGTACCAAGCCTTGAG 849

QY 41 ProValGlyLysAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 Db 850 CCGTGGGGGTCCGAGTCTGGGACAAAGATGCGACCAAGGACGCGATACAGGCCAGAG 909
 QY 61 TyrGluAspGlyArgGlyPheGlyLLeGlyGlnLeuValTrpGlyLysLeuArgGlyPhe 80
 Db 910 TACGAGAGCGCGCGGGCTTTGGCATTGGGGAGCTGTGTGGGGGAAACTGGCGGGCTTC 969
 QY 81 SerTrpTrpProGlyArgLLeValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 Db 970 TCTGTGGTGGCCAGCGCGCATTTGTCTGTGGATGACAGCGCGCGGAGCCGAGCAGCTGAA 1029
 QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGlnLys 120
 Db 1030 GGACCCCGCTGGGTCAATGTGTTCGAGACGGCAAAATTCACGTGTGTGTGTGGAAAG 1089
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 Db 1090 CTGATGCCGCTGAGCTCGTTTGCAGTGGCTTCCACGAGGCCACGTACAAACAGCAGCCC 1149
 QY 141 MetTyrArgLysAlaLLeTyrGlnValLLeGlnValAlaSerSerArgAlaGlyLysLeu 160
 Db 1150 ATGTACCGCAAGCCATCTACAGAGTCTGCAAGTGGCAGACGCGCGCGGGAAGCTG 1209
 QY 161 PheProAlaCysHisAspSerAspGlyLysAspSerGlyLysAlaValGlnAlaAsn 180
 Db 1210 TTCCCGGTGTGCGACAGCCAGTGAAGTGAACACTGCCAAGGCCGTGAGTGCAGAAC 1269
 QY 181 LysGlnMetLLeGlnTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGln 200
 Db 1270 AAGCCCATGATTGAATGAGGCGCTTGGGGGCTTCCAGCCTTTCGCGCTTAAGGCGCTGAG 1329
 QY 201 ProProGlnGlnGlnLysAsnProTyrLysGlnValTyrThrAspMetTrpValGluPro 220
 Db 1330 CCACCAGAAAGAGAGAGATCCCTACAAAGAGTACACAGACATGTGGGTGAACCT 1389
 QY 221 GlnAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
 Db 1390 GAGGACGCTGCTTACGACACCACTCCACAGCCAAAGACCCCGGAAGACACAGCGGAG 1449
 QY 241 LysProLysValLysGlnLLeLLeAspGlnArgThrArgGlnArgLeuValTyrGlnVal 260
 Db 1450 AAGCCCAAGTCAAGAGATTAATTGATGAGCCACACAGACGCGCTGTGTGACGAGTG 1509
 QY 261 ArgGlnLysCysArgAsnLLeGlnAspLLeCysLLeSerCysGlySerLeuAsnValThr 280
 Db 1510 CCGCAGAAAGTGCAGGAACATTGAGGACATCTGCATCTCTGTGGAGGCTCAATGTTACC 1569
 QY 281 LeuGlnHisProLeuPheLLeGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 Db 1570 CTGGACACCCCTCTCTTGTGGAGGATGTGCCAAAATCGCAAGAACTGCTTTCGGAG 1629
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrLLeCysCysGlyGly 320
 Db 1630 TGTGGCTACAGTACGACAGCAGCAGCGCTACAGTCTTACGACCACTGTGTGTGGGGG 1689
 QY 321 ArgGlnValLeuMetCysGlyLysAsnAsnAsnCysArgCysPheCysValGlnCysVal 340
 Db 1690 CGTGAAGTCTCATGTGCGGAACAAACAACCTCTGAGGTGCTTTGGCTGTGAGTGTG 1749
 QY 341 AspLeuLeuValGlyProGlyLLeAlaGlnAlaAlaLLeGlyGlnAspProTyrAsnCys 360
 Db 1750 GACCTCTTGTGTGGGCGCGGGCTGCCAGGACCATTAAGGAAGACCCCTGGAACCTG 1809
 QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGlnAspTrpProSer 380
 Db 1810 TACATGTGGGGCACAGAGGTACTTACGGGTGTGTGCGCGCGGAGAGAGACTGGCCCTCC 1869
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGlnPheAspProProLysValTyr 400
 Db 1870 CCGCTTCAGATGTTCTTCTGCTAATTAACACAGACAGGAATTTGACCTTCGAAAGCTTAC 1929

```

Oy 401 ProProValProIaGluValSerArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 1930 CCACCTGCTCCACCTTGAAGAAGAGAACCCATCCGGCTGCTCTCTCTTGAATGAAATC 1989
Oy 421 AIAthrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAspSer 440
Db 1990 GCTACAGAGGCTCTGGTGGCTGAAGAGACTTGGCATTGAGGTGAGCCGCTACATTGGCTCG 2049
Oy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 2050 GAGGTGTGTGAGGACTCCATCAGCGTGGCATGTGTGGGACCAAGGAGGAGATCATGTAC 2109
Oy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGlnIleTyrPheProPheAspLeu 480
Db 2110 GTGGGGAGAGTCCGACAGCTGACACAGAGCATATCCAGAGTGGGGCCATTGCAATCTG 2169
Oy 481 ValIleGlyIleSerProCysAsnAspLeuSerIleValAsnProIleArgLysGlyLeu 500
Db 2170 GTGATTGGGGGAGTCCCTGCAATGACCTCTCATGCACTGCAACCTGCTGCAAGGGCCTC 2229
Oy 501 Tyr-----GluGlyThrGlyArg 506
Db 2230 TACGGTAGACAGCCCGACGTGATGCTTCTTCGACCTCTCAGAGGGCAGCTGCGCCG 2289
Oy 507 LeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArg 526
Db 2290 CTCTTCTTGAAGTTCTACCGCTCTCTGATGATGCGCGCCCAAGAGGAGATGATCGC 2349
Oy 527 ProPhePheTyrLeuPheGluLeuValAlaIleMetGlyValSerAspLysArgAspIle 546
Db 2350 CCCTTCTTCTGGCTCTTTGGAATGTGTGGCCATGTGGGCTTGTGACAAAGAGGAGCATC 2409
Oy 547 SerArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaIleHis 566
Db 2410 TCGCAATTTCTCGAGTCCAAACCTGTGATGATTGATGACAAAGATGTGACGTGCACAC 2469
Oy 567 ArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrVal 586
Db 2470 AGGGCCCGTACTTCTGGGGTAACTTCCCGGTATGAAACAGGCCGTTGGCATCTGTG 2529
Oy 587 AsnAspLysLeuGluLeuGlnGlyCysLeuGluHisGlyArgIleAlaLysPheSerLys 606
Db 2530 AATGATTAACCTGAGCTGAGAGGTCTGTGAGACATGCGAGGTACCAAGTTCAACAAA 2589
Oy 607 ValAlaGlyThrIleThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPhePro 626
Db 2590 GTGAGAGACCATTAACGAGGTCAAACTCCATTAACAGAGGCAAAAGACAGCATTTTCT 2649
Oy 627 ValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlyMetGluArgValPheGly 646
Db 2650 GTCTTCATGAATGAGAAAGAGACATCTTATGTGTGACATGAAATGAAAGGATTTGGT 2709
Oy 647 PheProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuLeu 666
Db 2710 TTCCAGAGTCACTATCTACGCTCAACATGAGCCGCTTGGCGAGCAGACAGCTGTG 2769
Oy 667 GlyArgSerTyrSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPhe 686
Db 2770 GGGCCGATATGAGAGGTGCAGTCAATCCGCACTTGTGCTCGCTGAAGAGATATTTT 2829
Oy 687 AlaCysVal 689
Db 2830 CGCTGTGTG 2838

```

```

ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A., Yokoyama, S.,
Ohara, O., Nagase, T. and F. Kikuno, R.
TITLE None title
JOURNAL Published Only in Database (2005)
AUTHORS 2 (bases 1 to 4476)
TITLE Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A., Yokoyama, S.,
Ohara, O., Nagase, T. and F. Kikuno, R.
JOURNAL Direct Submission
COMMENT Submitted (12-MAR-2005) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research, 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail: cdhainfo@kazusa.or.jp,
URL: http://protein.gsc.riken.go.jp/, Tel: 81-438-52-3930,
Fax: 81-438-52-3931)
This work was supported in part by the National Project on Protein
Structural and Functional Analysis, Ministry of Education, Culture
, Sports, Science and Technology of Japan. Totoki Y., Toyoda A.,
Takeda T., Sakaki Y., Tanaka A., Yokoyama S., RIKEN Genomic Sciences
Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.
e-mail: aktanaka@postman.riken.go.jp
URL: http://protein.gsc.riken.go.jp/
Location/Qualifiers
1..4476
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pJ01914"
/tissue_type="brain"
/notes="this clone is also named as hsk003000130
vector: pBluescriptII SK plus"
<1..4476
/genes="DNA cytosine methyltransferase 3 alpha isoform a
variant"
<148..2583
/genes="DNA cytosine methyltransferase 3 alpha isoform a
variant"
/inference="non-experimental evidence, no additional
details recorded"
/notes="start codon is not identified."
/codon_start=1
/product="DNA cytosine methyltransferase 3 alpha isoform a
variant"
/protein_id="BAD92070.1"
/db_xref="GI:62087246"
/translation="LSPPKTTIIPSPKSSPPAPRSPSLPFGAOWPAMPSSGPGDTS
SAAEERERKQEEBEPKREPERSTLAKVRPGKRGHPVESDTPKDPVI
SKSPKADDSASELIPNGDLKREPERPEGSPAGCGGAPAGEGAETLPASR
AVENCCITPKESRGAPARAGKEQKTNIESMKMESRGRLRGGLHBESLRRPMPRL
TFQADPPIYISRRKDEWLARKRAEKAKYIAQMAVVENQGFGESEKVEASPPA
VOQPDPAIPVATIPPEVGSADGDKNATKADDEVEYDGFGLGLVWGKLGFS
WMPGRITVMMPPKRAEGRLVWVMPDGKSVVCEVLELMPISFCSFHQATYKQ
PMYRAIYEVLOVASRAKGLFPVCHDSEPTLAAVEONKPMIEMALGGPQSPK
GLEPPEKKNPKYETITPMVPEPAAAPAPKPKRSTKTEKXKVEIIEPRER
LVEVRQKRNIEDICISGLNVTLEHPLFYVGNCKQKCFLECAIYDDGTOSY
CTICCGEGLVLCGNNCCRCFVECVLIVPGAQAIAKEDPNNCYMGKGTGL
LRREDWPSRLQMPFANNHDEFPDPKYPAPKARPIRIYLSFDGLATGLVLKD
LGIQVDRYIASVECDSTVGWRIHQGLKMYGVDRSVTKQKIQWGPEDLIVGSPC
NDLSIVNARKGLVGTGRLPFEFVRLHLDARPKGDPRPFVLPENVVAMGVSDRD
ISRLFEVNSHLY"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0 Length: 4476
Score: 3653.00 Matches: 679
Percent Similarity: 90.1% Conservative: 2
Best Local Similarity: 89.8% Mismatches: 8
Query Match: 97.2% Indels: 67
DB: 5 Gaps: 1

```

```

RESULT 15
AB208833
LOCUS AB208833 4476 bp mRNA linear PRI 31-MAR-2005
DEFINITION Homo sapiens mRNA for DNA cytosine methyltransferase 3 alpha
isoform a variant protein.
ACCESSION AB208833
VERSION AB208833.1 GI:62087245
KEYWORDS FLI CDNA.
SOURCE Homo sapiens (human)

```

US-10-623-813-85 (1-688) x AB208833 (1-4476)
 QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnLysValGluGluAlaSer 20
 Db 907 ATGAATGCTGTGGAAAGAAAACGAGGGGCGGGGAGCTCAGAAAGTGAGAGGAGCGAGC 966
 QY 21 ProProAlaValAlaGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 Db 967 CCTTCCTGCTGTGACGAGCCCACTGACCCCGCATCCCTGCTGGCTTACCGAGCTGAG 1026
 QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 Db 1027 CCGGTGGGGTCCCATGCTGGGGACAAAGATGCCCAAGACGGCGATGACGAGCCAG 1086
 QY 61 TyrGluAspGlyArgGlyPheGlyLeuValTyrGlyLysLeuArgGlyPhe 80
 Db 1087 TAGAGAGACGGCGGGGCTTTGGCATTTGGGAGCTGTGTGGGAAATCGGGGCTTC 1146
 QY 81 SerTrpTrpProGlyValGlyLeuAlaSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 Db 1147 TCTGTGGGCGCAGGCGCGCATTTGTCTTGTGTGATGACGGGCGGAGCCGAGCGCTGAA 1206
 QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 Db 1207 GGCACCGCTGGGTCAATGCTTGGAGACGGCAAATTCAGTGTGTGTGTGAGAG 1266
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 Db 1267 CTGATGCGCGTGAGCTGTTTTTGCAGTGCCTTCCACGAGGCCAGTTCACAAAGCAGCCCC 1326
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 Db 1327 ATGTACCGCAAGACCATCTACGAGTCTCTGAGTGCACGAGCCCGGGGGAACTG 1386
 QY 161 PheProAlaCysHisAspSerAspGlySerAspSerGlyLysAlaValGluValGlnAsn 180
 Db 1387 TTCCCGGTGCGCACGACGAGATGAGAGTGAAGTCACTCCAAAGCCGATGAGGTGCAGAAC 1446
 QY 181 LysGlnMetIleGluTTPAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 Db 1447 AAGCCCATGATTGATTGAGGCGCTGGGGGCTTCCAGGCTTCTGCTTCAAGGGCGCTGAG 1506
 QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTTPValGluPro 220
 Db 1507 CCACCAAGAAAGAAAGAAATCTCTACAAAGATGTACAGGACATGTGGTGAACCT 1566
 QY 221 GlnAlaAlaAlaTyrAlaProProProAlaLysLysProArgLysSerThrThrGlu 240
 Db 1567 GAGGCACTGCTTACGACCACTCCACGACCAAAAGCCCGGAAGAGACAGCGGAG 1626
 QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
 Db 1627 AAGCCCAAGGTCAAGAGATTATTGATGAGGCAAGAGAGCGCGTGTGTACGAGGTG 1686
 QY 261 ArgGlnLysCysArgAsnIleGluAspLecCysIleSerCysGlySerLeuAsnValThr 280
 Db 1687 CGCAGAAAGTCCGGAACATTGAGACATCTGATCTCTGTGGAACCTCAATATTGAACC 1746
 QY 281 LeuGlnHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 Db 1747 CTGGAACACCCCTCTTCTGTGGAGATGTGCCAAACCTGCAGAACTGCTTCTGAG 1806
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 Db 1807 TGTGCGTACAGTACGACGACGAGCTACCACTCACTGACCACTGTGTGGGGGC 1866
 QY 321 ArgGlnValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 Db 1867 CGTAGAGTGTCTCATGTGGGAAACAACATCTGCTGAGGTCTTTTGTGGAGATGTGTG 1926
 QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTyrAsnCys 360
 Db 1927 GACTCTTGTGTGGGGCGGGGCTGCCAGGACCATTTAAGAAAGCCCTTGAACCTGC 1986

QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
 Db 1987 TACATGTGGGGACAAAGGTACTTACGGGCTGTGCGGGGAGAGAGACTGGCCCTTC 2046
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 Db 2047 CCGCTCCAGATGTCTTCTGCTTAATAACACGACGACGAGAAATTTGACCTCCAAAGTTTAC 2106
 QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 Db 2107 CCACCTGTCCCAAGTGAAGAGAAAGCCCATCCGGGTCTGTCTCTTTGATGGAATC 2166
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 Db 2167 GCTACAGGGCTCTGTGTGTGAAGACTTGGGCAATTCAGGTGACCCCTACATGTGCTTCG 2226
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 Db 2227 GAGGTGTGAGGACTCCATCACGGTGGCATGTGTGGCACAGGGGAAGATCATGTAC 2286
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
 Db 2287 GTCGGGGACGTCCGCGAGCTCACAGAAACATATCAGAGAGTGGGCCCATTCGATCTG 2346
 QY 481 ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 Db 2347 GTGATTTGGGGAGAGTCTCTGCAATGACTTCTCATGTCAACCTCTCGCAAGGGCTC 2406
 QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 Db 2407 TACGAGGGACGTGGCGGCTCTTCTTGAATTCTACCGCTCTGTGATGTGGCGGGCC 2466
 QY 521 LysGluGlyAspAspArgProPhePheThrPhePheGluAsnValValAlaMetGlyVal 540
 Db 2467 AAGGAGGAGATGATGCCCTTCTTGTGGAAGTGTGTGTCATGAGGCGGT 2526
 QY 541 SerAspLysArgAspIleSerArgPheLeu----- 550
 Db 2527 AGTCAAGAGGAGCATCTCGCATTTCTCGAGGTCAAAATAGCAACCTGTACTAGTG 2586
 QY 550 ----- 550
 Db 2587 GCTAGTGGCACTGAATTGAACAGTGTAGAGCAACCAACCATCCGAGAGTTCAATTG 2646
 QY 550 ----- 550
 Db 2647 GACATGTGATCTAGAACGTTCCAGTGAACGTTGTGAATGTGGCTAGGTAAATTCC 2706
 QY 551 -----GluSerAsnPr 554
 Db 2707 AGCTTTCTCGAGGCTGCAGGCTAGCCAGTGTGGCTCTCGAAGAGAAATCCAACC 2766
 QY 554 OValMetIleAspAlaLysGluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAs 574
 Db 2767 TGTGATGATTGAAGCCAAAGAGTGTCACTGACACAGAGGCGCGCTACTTCTGGGTAA 2826
 QY 574 nLeuProGlyMetAsnArgProLeuAlaSerThrValAsnAspLysLeuGlnLeuGlnG 594
 Db 2827 CCTTCCCGGATGAAACAGGCCGCTTGGCATCTGTAATGATTAACCTGAGCTGACGGA 2886
 QY 594 uCysLeuGlnHisGlyArgIleAlaLysPheSerLysValArgThrIleThrThrArgSe 614
 Db 2887 GTGTCTGGAGCATGGACGATACCAAGTTACCAAAAGTGAAGCAATTAACACGAGTGC 2946
 QY 614 rAsnSerIleLysGlnGlyLysAspGlnHisPheProValPheMetAsnGlnLysGluAs 634
 Db 2947 AAACCTCATTAAGCAAGGCAAGAGCAGCATTTTCTGTCTTCAATAATAGAAAGGA 3006
 QY 634 pIleLeuTrpCysThrGluMetGluArgValPheGlyPheProValHisTyrThrAspVa 654
 Db 3007 CATCTTAGTGTGACATGAATGAAAGGATTTGTGTTCCAGTCCACTATCTAGAGCT 3066

Qy	654	1SerAsnMetSerArgLeuAlaArgGlnArgLeuLeuGlyArgSerTrpSerValProVa	674
Db	3067	CTCCAACTATGAGCCGCTTTGGGAGGCAAGACTGCTGGCCGGTCAATGAGCGTGCAGT	3126
Qy	674	111eArgHisLeuPheAlaProLeuLysGluTyrPheAlaCysVal	689
Db	3127	CATCGGCACCTTTGCGCTCGCTGAAGAGTATTTTGGGTGTGTG	3172

Search completed: October 10, 2006, 17:31:30
Job time : 10584 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 10, 2006, 14:46:24 : Search time 2084 Seconds
(without alignments)
6093.700 Million cell updates/sec

Title: US-10-623-813-85

Perfect score: 3758
Sequence: 1 MNAVENQASGSEQKVEAS.....WSVPVIRHLFADLKEYFACV 689

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+gdn.model -DEV=x1h
-O=/abs/ABSWEB.sppol/US10623813/runat_10102006.143521.7265/app_query.fasta_1
-DB=Published Applications NA.Main -QFMT=fastap -SUFRFX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR.SCORE=pct
-THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MIMLEN=0 -MAXLEN=200000000 -HOST=abs03h
-USER=US10623813@CCN 1.1.1675@runat_10102006.143521.7265 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	3758	100.0	2318 9	US-10-623-813-83 Sequence 83, Appl

2	3758	100.0	4192	5	US-09-720-086-1	Sequence 1, Appl1
3	3758	100.0	4192	9	US-10-623-813-1	Sequence 1, Appl1
4	3758	100.0	6157	11	US-10-330-773-276	Sequence 276, App
5	3714	98.8	2371	9	US-10-623-813-84	Sequence 84, Appl1
6	3714	98.8	3005	6	US-10-144-577-2	Sequence 2, Appl1
7	3714	98.8	3005	6	US-10-172-118-132	Sequence 132, App
8	3714	98.8	3005	8	US-10-342-887-132	Sequence 132, App
9	3714	98.8	3009	11	US-10-330-773-279	Sequence 279, App
10	3714	98.8	4293	5	US-09-720-086-3	Sequence 3, Appl1
11	3714	98.8	4293	9	US-10-623-813-3	Sequence 3, Appl1
12	2708	72.1	2338	7	US-10-264-237-953	Sequence 953, App
13	2159.5	57.5	4145	5	US-09-720-086-4	Sequence 4, Appl1
14	2159.5	57.5	4145	6	US-10-144-577-3	Sequence 3, Appl1
15	2159.5	57.5	4145	6	US-10-144-577-38	Sequence 38, Appl1
16	2159.5	57.5	4145	7	US-10-172-118-1337	Sequence 1337, App
17	2159.5	57.5	4145	8	US-10-342-887-1337	Sequence 1337, App
18	2159.5	57.5	4145	9	US-10-623-813-4	Sequence 4, Appl1
19	2159.5	57.5	4195	5	US-09-720-086-2	Sequence 2, Appl1
20	2159.5	57.5	4195	9	US-10-623-813-2	Sequence 2, Appl1
21	1881	50.1	3897	6	US-10-144-577-39	Sequence 39, Appl1
22	1837.5	48.9	95484	11	US-10-330-773-275	Sequence 275, App
23	1722.5	45.8	127722	11	US-10-330-773-278	Sequence 278, App
24	1318	35.1	1060	6	US-10-106-698-926	Sequence 926, App
25	1095.5	29.2	709	3	US-09-969-034-1235	Sequence 1235, App
26	1003	26.7	711	10	US-10-450-763-21969	Sequence 21969, A
27	944	25.1	622	12	US-10-301-480-606665	Sequence 606665
28	942	25.1	994	12	US-10-301-480-606665	Sequence 120074
29	942	25.1	994	12	US-10-301-480-606665	Sequence 120074
30	872.5	23.2	2057	6	US-10-144-577-40	Sequence 40, Appl1
31	785	20.9 <td>2008</td> <td>6</td> <td>US-10-144-577-41</td> <td>Sequence 41, Appl1</td>	2008	6	US-10-144-577-41	Sequence 41, Appl1
32	750	20.0	440	9	US-09-720-086-54	Sequence 54, Appl1
33	750	20.0	440	9	US-10-623-813-54	Sequence 54, Appl1
34	688	18.3 <td>1705</td> <td>8</td> <td>US-10-479-334-35</td> <td>Sequence 35, Appl1</td>	1705	8	US-10-479-334-35	Sequence 35, Appl1
35	685	18.2 <td>2951</td> <td>10</td> <td>US-10-450-763-21970</td> <td>Sequence 21970, A</td>	2951	10	US-10-450-763-21970	Sequence 21970, A
36	644.5	17.2	1670	8	US-10-479-334-21	Sequence 21, Appl1
37	639	17.0	1521	8	US-10-479-334-7	Sequence 7, Appl1
38	636	16.9	522	5	US-09-720-086-61	Sequence 61, Appl1
39	636	16.9	522	6	US-10-623-813-61	Sequence 61, Appl1
40	629	16.7	598	6	US-10-027-633-199022	Sequence 199022
41	629	16.7	598	6	US-10-027-633-199022	Sequence 199022
42	629	16.7	598	6	US-10-027-633-199022	Sequence 199022
43	629	16.7	598	7	US-10-027-633-199023	Sequence 199023
44	609	16.2	435	5	US-09-720-086-18	Sequence 18, Appl1
45	609	16.2	435	9	US-10-623-813-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-10-623-813-83
; Sequence 83, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
; APPLICANT: LI, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; APPLICANT: Chen, Taiping
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Use
; FILE REFERENCE: 0609.4560003
; CURRENT FILING DATE: US/10/623,813
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 09/720,086
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: PCT/US99/14373
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,906
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 60/093,993
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83

```

; LENGTH: 2318
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-623-813-83

Alignment Scores:
Pred. No.: 0 Length: 2318
Score: 3758.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0

US-10-623-813-85 (1-689) x US-10-623-813-83 (1-2318)
Oy 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnValGluGlnAlaSer 20
Db 148 ATGAAATGCTGTGAAAGAAACCAAGCCTCTGAGAGTCTCAAGAGTGGAGAGGCCAGC 207
Oy 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db 208 CCTCCTGCTGTGACGACGCCACGAGACCTGCTTCCGACTGTGGCCACACCCCTGAG 267
Oy 41 ProValGlyGlyAspAlaGlyAspIleAsnAlaThrIleValAlaAspAspGluProGlu 60
Db 268 CCGATGAGAGGAGGATGCTGGGGACAAAGATGCTCAAGACAGCCGACGATGAGCTGAG 327
Oy 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValThrGlyIleuAspArgGlyPhe 80
Db 328 TATGAGAGTGGCCGGGCTTTGGCATTGAGAGCTGCTGTGGGGAAACTTGGGGCTTC 387
Oy 81 SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyIleArgSerArgAlaIleGlu 100
Db 388 TCTGTGTGGCCAGCCGCAATGTGTCTTGTGTGATGACAGCCGAGCCGAGACGCTGAA 447
Oy 101 GlyThrArgThrAlaMetTrpPheGlyAspGlyIleAspSerValValCysValGluIle 120
Db 448 GAGCATGCTGGGTCATGTGTTGAGATGGCAAGTCTTCAAGTGTGTGTGGAGAG 507
Oy 121 LeuMetProLeuSerSerPheCysSerAlaPheIleGlnAlaThrTyrAsnIleGlnPro 140
Db 508 CTGATCCCGTGAAGCTCTTCTGAGTGCATTCCACAGGCCACTTACAAAGCAGCC 567
Oy 141 MetTyrArgIleAlaIleTyrGluValIleuGlnValAlaSerSerArgAlaGlyIleu 160
Db 568 ATCTACCGCAAGCATCTCAAGATGCTCAGGTGGCCAGCCCTGCGGGAACCTG 627
Oy 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyIleValAlaGluValIleGln 180
Db 628 TTTCGAGCTTGCATGACATGATGAAGATGACATGGCAAGGCTGTGGAAATGCGAAC 687
Oy 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProIleGlyLeuGlu 200
Db 688 AAGCAGATGATTGATGGCCCTCGTGGCTTCCAGCCCTCGGCTCTTAAGGGGCTGAG 747
Oy 201 ProProGluGluGluIleAsnProTyrIleGluValIleThrAspMetTrpValGluPro 220
Db 748 CCACCAAGAAAGAAAGAAATCCTTACAGAAATTTACACCAAGCAAGTGGGTGAGCCT 807
Oy 221 GluAlaIleAlaTyrAlaProProProProAlaIleIleIleIleIleIleIleIleIle 240
Db 808 GAAAGACCTGCTTACGCCCAACCCCAACCCCAAGCAAGAAACCAAGAAAGACACAGAG 867
Oy 241 LysProIleValIleGluIleIleAspGluArgThrArgIleuArgIleValIleGluVal 260
Db 868 AAGCCTAAGGTCAAGAGATCATTTGATGAGGCACAAGGGAGCGGCTGTATGAGGTG 927
Oy 261 ArgGlnIleCysArgAsnIleGluAspIleCysIleSerCysGlySerIleuAsnValThr 280
Db 928 CGCCAGAAAGTGCAGAAACATCGAGACATTTGATCTCATGTGGAGCCCTCAAGTCAAC 987
Oy 281 LeuGluIleProLeuPheIleGlyIleCysGlnIleAsnCysIleAsnCysPheLeuGlu 300

```

```

Db 988 CTGAGACACCACTTCTTCAATTGGAGCATGTGCCAAGACTGTAAGAACTGTTTGGAG 1047
Oy 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1048 TGTGCTTACAGATGACACACATGGGTACGATCTTATGACACCTTGTGTGGGGGG 1107
Oy 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1108 CCGAAGTCTCTCATGTGTGGAAACAACATCTGCGAGGTGCTTTGTGTCCAGTGTGTG 1167
Oy 341 AspLeuLeuValGlyProGlyAlaIleGlnAlaAlaIleLysGluAspProTyrAsnCys 360
Db 1168 GATCTCTTGTGGGGGCAAGAGCTGCTCAGGACGACATTAAAGAAAGACCCCTGGAAC 1227
Oy 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuAspArgArgGluAspTrpProSer 380
Db 1228 TACATGTGCGGCATTAAGGACCTATGTGGCTGTCCGAAGACGGGAACATGGCTTCT 1287
Oy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProIleValTyr 400
Db 1288 CAGCTCCAGATGTTCTTTGCCAAATTAACATGACAGAAATTTGACCCCAAGGTTTAC 1347
Oy 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 1348 CCACTGTGCGACGTGAGAAAGAAAGCCCATCCGGCTGCTGTCTTGTATGGGATT 1407
Oy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 1408 GCTACAGGGCTCTGTGTGTGAAGACCTGGGCATCCAAAGTGAACCGCTACATGCTCC 1467
Oy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyIleLysIleMetTyr 460
Db 1468 GAGGTGTGAGAGATCTCCATCACGGTGGGACATGTCGGCACAGGAAAGATCATGTAC 1527
Oy 461 ValGlyAspValArgSerValThrGlnHisIleGlnGluTrpGlyProPheAspLeu 480
Db 1528 GTCGGGGAGTCGCCGACGCTCACAGAGCATATCCAGATGGGGCCCATTCGACCTG 1587
Oy 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgIleGlyLeu 500
Db 1588 GTGATGGAGGAGATGCCCTCGCAATGACCTCTCATTTGAACCTGCGCGCAAGGACCT 1647
Oy 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 1648 TATGAGGATATGGCCGCTCTTCTTGAGTTCTACCGCTCTGATGATGGCGGCC 1707
Oy 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
Db 1708 AAGGAGGAGATGATGCGCCCTTCTGTGCTTGAAGAAATGTGTGGCCATGGCGCTT 1767
Oy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 1768 AGTGACAAAGGAGACATCTCGGATTTCTTGAGTCTTAACCCCGTGTATGATGCGCAAA 1827
Oy 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db 1828 GAAATGTCTGCTGCACACAGGCGCTTACTTGTGGGTAACCTTCTGGAATGAACAG 1887
Oy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGluCysLeuGluHisGlyArg 600
Db 1888 CTTTGGCATCCACTGTGATGATTAAGCTGAGCTGCCAAGATGCTGAGACCGCAACA 1947
Oy 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db 1948 ATAGCCAAAGTTCAAGAAATGAGACCATTTACCAAGGCTCAAACTTAAAGCAGGGC 2007
Oy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db 2008 AAAGACCAACATTTCCCGCTCTTCAAGACGAGAAAGAAAGAACTCTGTGTGTCATCTGA 2067
Oy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2068 ATGAAAGAGGTGTTGTGCTTCCCGCTCACTACAGAGATGCTTCCAACATGAGCCGCTTG 2127

```

QY 661 AAlaArgLnaTArgLeuLGLVArgSerTrpSerValProValIleArgHisLeuPheAla 680
 DB 2128 GCGAGGCGAGAGACTGCTGGGCGGATCTGGAGCGTGGCGGATCTCCGACCTCTTGCT 2187
 QY 681 ProLeuLysGluTrpPheAlaCysVal 689
 DB 2188 CCGCTGAAGGAATATTTTGTGTTGTG 2214
 RESULT 2
 US-09-720-086-1
 ; Sequence 1, Application US/09720086
 ; Publication No. US20060084053A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, En
 ; APPLICANT: Okano, Masaki
 ; APPLICANT: Xie, Shaoping
 ; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
 ; TITLE OF INVENTION: Polypeptides & Uses Thereof
 ; FILE REFERENCE: 0609, 4560002
 ; CURRENT APPLICATION NUMBER: US/09/720, 086
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/14373
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/090,906
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/093,993
 ; PRIOR FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 4192
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (4161)..(4161)
 ; OTHER INFORMATION: May be any nucleic acid
 US-09-720-086-1
 Alignment Scores:
 Pred. No.: 0 Length: 4192
 Score: 3758.00 Matches: 689
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Gaps: 0
 DB: 5
 US-10-623-813-85 (1-689) x US-09-720-086-1 (1-4192)
 QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnLysValGluGluAlaSer 20
 DB 874 ATGAAATGCTGTGAAGAAGCAAGGCTCTGGAGAGCTCAGAAAGTGGAGGAGGCGCAGC 933
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrTrpProGlu 40
 DB 934 CCTCTGCTGTGACGACGCCACGACCTGCTTCTCCGACTGCTGGCGACCACTCCCTGAG 993
 QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 DB 994 CCAAGTGAAGGAGGATGCTGGGGAACAAGATGCTACCAAGACACCGACGATGAGCTGAG 1053
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
 DB 1054 TAGAGAGATGGCGGGCTTTGGCATTGGAGAGCTGTGTGGGGAAACTTCGGGGCTTC 1113
 QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 DB 1114 TCTGTGTGGCGACGCGGAATGTGTCTGTGTGATGACAGCGCGAGCCGACGACGTGAA 1173
 QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 1174 GGCACTGCTGCTCATGTGTGTGCGAGATGCGCAAGTTCTCAGTGTGTGTGTGAGAAAG 1233

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGluAlaThrTrpAsnLysGlnPro 140
 DB 1234 CTATGCGCTGTAGCTCTTCTTGACAGTTCATCCACAGGCCCACTTCAACAAAGACAGCCC 1293
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 1294 ATGTACCGCAAGAACCATTTACGAAGTCTTCAGGTGGCCAGCGCCGCGGAAAGCTG 1353
 QY 161 PheProAlaCysHisAspSerAspGluSerSerGlyLysAlaValGluValGlnAsn 180
 DB 1354 TTTCAGCTTGGCCATATACAGTGTGAAGTACAGATGGCAAGGCTGTGAAGTGCAGAAC 1413
 QY 181 LysGluMetIleGluTrpAlaLeuGlyGlyPheGluProSerGlyProLysGlyLeuGlu 200
 DB 1414 AACCAATGATTAATGAGGCCCTCGGTGCTTCAAGCCCTCGGCTCTTAAGGCTCGAG 1473
 QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
 DB 1474 CCACGAAAGAAAGAAAGAAATCCTTCAAGAAAGTTTACCCGACATGTGTGTGAGCCT 1533
 QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrTrpGlu 240
 DB 1534 GAAGCAGCTGCTTACCCGCCACCCACCCACGCAAGAAACCCAGAAAGAGACACAGAG 1593
 QY 241 LysProLysValLysGluIleLeuAspGluArgTrpArgGluArgLysValTyrGluVal 260
 DB 1594 AAACCTTAAGGTCAAGAGATCATTGATGAGCGCAAGAGGAGCGGCTGTGTGTAGGTG 1653
 QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1654 CGCCAAAGTGAAGAAATCCAGAGCATTTGATCTCATGTGGAGCCCAATGTGCACC 1713
 QY 281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 DB 1714 CTGGACACCCCACTCTTCAATTGAGGCAATGTCACAACTGTAAAGACTCTTGTGGAG 1773
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGly 320
 DB 1774 TGTCTTACAGATGACACAGATGGGTACAGTCTTATTTGACCACTGTGCTGTGGGGG 1833
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 DB 1834 CGTGAAGTCTCATGTGTGGAAACAACACTGTGAGGTCCTTTGTGCGAGTGTGTG 1893
 QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
 DB 1894 GATCTCTTGTGGGCGCAAGAGCTGCTCAGGCAAGCATTAAGGAAGAACCCCTGGAAC 1953
 QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLysLeuArgArgGluAspTrpProSer 380
 DB 1954 TACATGTGGGCAATGAGGACCTATGAGGCTGTGCTGCAAGAGGGAAGACTGGCTTCT 2013
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 DB 2014 CCACTCAGATGTCTTGTCCAAATTAACATGACAGGAATTTGACCCCAAGGTTTAC 2073
 QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2074 CCACTGTGCGCAGCTGAGAAAGAAAGCCATCCGGTGTCTCTCTTTGATGAGGAA 2133
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 DB 2134 GCTACAGGCTCTGTGTGTGAAGACTGGGCAATCCAAAGTGAACGCTACATTGCTCTC 2193
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2194 GAGGTGTGAGAGACTCCATCAACGTTGGGCTGTGTGCGCACCAAGGAAAGATCATGTAC 2253
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
 DB 2254 GTGGGGAGGTCGAGGCTCACAGAAAGCATATTCAGAGATGGGGCCCATTTGCACCTG 2313

```

QY 481 vallllgllyserprocyasnaaspleuSerlleValaenproAlaarglysglyleu 500
DB 2314 GTGATGGAGGAGGAGTCCCTCAATGAGCTCTCAATGTCAACCTGCCCCCAAGGAGCTT 2373
QY 501 TTTGUGLlyThrglyArgleuPheGluPheThyTrargleuHissapaAargPro 520
DB 2374 TATGAGGGATCTGGCCGCTCTTTGAGTTCTACCCGCTCTCGCATGATGCCGCC 2433
QY 521 LysGluGlyAspaAspaArgProPhePheTrpleuPheGluasnValAlaIamGlyVal 540
DB 2434 AAGGAGGAGATGATGCCCTTCTTCTGCTCTTGAAGATGGTGGCCATGGCGCTT 2493
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerasnProValMetIleAspAlaLys 560
DB 2494 AGTGACAAAGAGGAGCATCTGCGCATTTCTTGAGCTAACCCCGGATGATGAGCCAAA 2553
QY 561 GluValSerAlaIaIaHissArgAlaArgTyrrPheTrpGlyAsnLeuProGlyMetAsnArg 580
DB 2554 GAAGTGTCTCTGCACACAGGGCCGCTTACTTCTGGGGTAACTTCTGGCATGACAG 2613
QY 581 ProLeuAlaSerThrValaAsnAspLysLeuGluLeuGluGlyCysLeuGluHissGlyArg 600
DB 2614 CTTTGGCATCCACTGTGAATGATTAAGCTGAGCTGCAAGATGTCTGGAGCACGCGAGA 2673
QY 601 IleAlaLysPheSerLysValArgThrIleThrTrpArgSerAsnSerIleLysGlnGly 620
DB 2674 ATAGCCAAAGTTCAGCAAGAGGACCAATTAACCCAGGTCAACTTAATAAGCAGGCG 2733
QY 621 LysAspGlnHissPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
DB 2734 AAAGACACAGATTTCCCGCTTTCATACAGAGAGAGAGACATCTCTGTGTGACCTGAA 2793
QY 641 MetGluArgValPheGlyPheProValHissTyrrThraspValSerAsnMetSerArgLeu 660
DB 2794 ATGAAAGGGTGTGGCTTCCCGCTTCCATACAGACAGCGTCTCAACATGAGCGCGCTTG 2853
QY 661 AlaArgGluArgLeuLeuGlyArgSerTrpSerValProValIleArgHissLeuPheAla 680
DB 2854 GCAGAGCAGAGACTGCTGGGCCGAGTCGTGAGCGTGGCGATCCGACCTCTTGTGCT 2913
QY 681 ProLeuLysGluTyrrPheAlaCysVal 689
DB 2914 CCGCTGAGAGGAATATTTTGTGTGTG 2940

RESULT 3
US-10-623-813-1
; Sequence 1, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; APPLICANT: Chen, Taiping
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
; TITLE OF INVENTION: Theoreof
; FILE REFERENCE: 0609, 4560003
; CURRENT APPLICATION NUMBER: US/10/623, 813
; PRIOR APPLICATION NUMBER: 2003-07-22
; PRIOR FILING DATE: 1999-06-25
; PRIOR FILING DATE: 1999-06-25
; PRIOR FILING DATE: 1999-06-25
; PRIOR FILING DATE: 1998-06-25
; PRIOR FILING DATE: 1998-06-25
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 4192
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (4161)..(4161)
; OTHER INFORMATION: n is a, c, g, or t
US-10-623-813-1

Alignment Scores:
Pred. No.: 0
Score: 3758.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 9
Gaps: 0

US-10-623-813-85 (1-689) x US-10-623-813-1 (1-4192)

QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnLysValGluGluAlaSer 20
DB 874 ATGAATGCTGTGAAAGAAACCAAGCGCTCTGGAGATCTCAGAAAGCTGGAGAGCCAGC 933
QY 21 ProProAlaValGlnGlnProThraAspProAlaSerProThrValAlaThrTrpProGlu 40
DB 934 CTTCTGCTGTGACAGAGCCCAAGCCCTGCTTCTCGACTGTGGCCACACCCCTTGAG 993
QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
DB 994 CAGTGAAGAGGGATGCTGGGGGCAAGAAATGTACCAAGCAAGCCGACATGAGCTGAG 1053
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValITrpglyLysLeuArgGlyPhe 80
DB 1054 TATGAGAGATGGCGCGGGCTTTGGCATTGAGAGCTGTGTGGGGAACCTTGGGGCTTC 1113
QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaIaGlu 100
DB 1114 TCCTGTGTGGCAGGCGCAATGTGTCTTGTGTGTGATACAGCGCGAGCCGAGCAGCTGAA 1173
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValCysValGluLys 120
DB 1174 GGCATCGCTGGGTATGATGTGTTCGAGATGGCAAGTTCTCAATGTGTGTGTGGAGAG 1233
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHissGlnAlaThrTyrrAsnLysGlnPro 140
DB 1234 CTCATCCCGCTGAGCTCTTCTGTGACGTGATTCACAGGCGCACTTAACAAGAGGCC 1293
QY 141 MetTyrrArgLysAlaIleTyrrGluValLeuGluValAlaSerSerArgAlaGlyLysLeu 160
DB 1294 ATGTACCGCAAGCCATCTACGAATCTCTCAGAGTGGCCAGCAGCGCTGCGGGAAGCTG 1353
QY 161 PheProAlaCysHissAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180
DB 1354 TTTCAGCTTGGCATGACAGTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1413
QY 181 LysGluMetIleGluITrpaAlaLeuGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 1414 AAGCAGATGATTTGAAGAGGCGCTCGGTGCTTCCAGCCCTCGGGTCTTAAGGGCTGGAG 1473
QY 201 ProProGluGluGluLysAsnProTyrrLysGluValTyrrThraspMetTrpValGluPro 220
DB 1474 CACCGAAGAGAGAGAGAAATCTTCAAGAGAGTTTACACCGACATGTGTGGAGGCT 1533
QY 221 GluAlaAlaAlaTyrrAlaProProProProAlaLysLysProArgLysSerThrTrpGlu 240
DB 1534 GAAGAGAGCTTACCCGCCACCCACAGCCCAAGAAACCCAGAAAGAGACACAGAG 1593
QY 241 LysProLysValLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 260
DB 1594 AAACCTTAAGTCAAGAGATCATTAATGAGCGCAAGAGAGAGAGAGAGAGAGAGAG 1653
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1654 CGCCAGAGTGCAGAAACATCGAGGACATTTGTATCTCATGTGGAGCCCAATGTGCACC 1713
QY 281 LeuGluHissProLeuPheIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 300

```

```

Db      1714 CTGAGACACCACTTTCATTGGAGGCAATGCGCAAGTCTTGAAGTCTTGGAG 1773
Qy      301 CysAlaTyrgLThrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1774 TGTGCTTACAGATATACAGACGATGGGTACCACTCTATTGCAACATCTGTGTGGGGG 1833
Qy      321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGlyCysVal 340
Db      1834 CGTGAAGTCTCATGTCTGGGAACAACTGCTGCAGAGCTTTTGTGTGCGATGTGTG 1893
Qy      341 AspleuLeuValGlyProGlyAlaAlaGlnAlaAlaIleGlyGlnAspProTyrPancCys 360
Db      1894 GATCTCTTGTGGGCGCAGAGAGCTGCTCAGGAGCCATTAGAGAAACCCCTTGAAGTCTGC 1953
Qy      361 TyrMetCysGlyAlaGlyGlyTyrTyrGlyLeuLeuArgArgGlyLysAspTyrProSer 380
Db      1954 TACATGTGGGGCATAGGGCACTTATGGCTGCTCGAAGAGGGAGAGCTGGCTTCT 2013
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProTyrValTyr 400
Db      2014 CCACTCCAGATGTTCTTTCGCAATACCATGACCAAGAAATTGACCCCCCAAGTTTAC 2073
Qy      401 ProProValProAlaGlnLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2074 CCACCTGTCCAGCTGAGAAAGAGAGAGCCCATCCGGTGTCTCTCTTGTATGGAGATT 2133
Qy      421 AlaThrGlyLeuLeuValLeuLysAspleuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2134 GCTACAGGGCTCTGTGTCTGAAGACCTGGGCATCAAGTGACCGTACATTTGGCTCC 2193
Qy      441 GluValCysGlnAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2194 GAGGTGTGAGGACTCCATCAAGGTGGGATGTGTGCGGACCAAGGAAAGATCATCTAC 2253
Qy      461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrPhePheAspleu 480
Db      2254 GTCGGGGAGCTCCGACGCTCACAGAACATATACAGAGAGGGGCCCATTCGACCTG 2313
Qy      481 ValIleGlyGlySerProCysAsnAspleuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2314 GTGATTGGAGGAGCTCCCTGCAATGACCTTCCATTGTCAACCTCCCGCAAGGAGACTT 2373
Qy      501 TyrGlnGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      2374 TATGAGGAGTACGTGGCGCTCTTCTTGAATCTTACCGCTCTTGATATGGCGGCCCC 2433
Qy      521 LysGlnGlyAspAspArgProPhePheTyrLeuPheGluAsnValValAlaMetGlyVal 540
Db      2434 AAGGAGGAGATGATGCGCCCTTCTTCTGGCTTGTGAGATGTGTGGCATGGCGGCTT 2493
Qy      541 SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys 560
Db      2494 AGTGACAAAGGAGACATCTCGCAATTTCTTGAGTCTAACCCCGTGATGATGCGCAAA 2553
Qy      561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
Db      2554 GAAAGTCTCTGCTGCACACAGGCGCTTACTTGTGGGTAACTTCTGCGCAAGAACAG 2613
Qy      581 ProLeuAlaSerThrValAsnAspLysLeuGlnGlnLysLeuGlnHisGlyArg 600
Db      2614 CCTTTGGCATCCACTGTGATGATTAAGCTGAGACTCAAGAGTGTCTGGAGCAGCGCAAA 2673
Qy      601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2674 ATAGCCAAAGTTACAGCAAGTGAAGACATTACACACAGGTCAAACTCTTAAAGCAGGGC 2733
Qy      621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTyrCysThrGlu 640
Db      2734 AAAGACCAAGCATTTCCCGCTTTCATGAACGAGAAAGAGACATCTGTGTGTCACTGAA 2793
Qy      641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2794 ATGGAAAGGGTGTGTGGCTTCCCGTCCACTACACAGAGCTGTCCAACATGAAGCCCTTG 2853

```

```

Qy      661 AlaArgGlnArgLeuGlnGlyArgSerTyrSerValProValIleArgHisLeuPheAla 680
Db      2854 GCGAGGCAAGACATCTGTGGCGCGGATCGTGAACGTCGCGGTATCGCCACCTTCGCT 2913
Qy      681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2914 CCGCTGAAGGAATATTTGCTTGTGTG 2940

RESULT 4
US-10-330-773-276
; Sequence 276, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 6157
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-330-773-276

Alignment Scores:
Pred. No.: 0 Length: 6157
Score: 3758.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 11 Gaps: 0

US-10-623-813-85 (1-689) x US-10-330-773-276 (1-6157)
Qy      1 MetAlaAlaValAlaGlnGlyLysAsnGlnAlaSerGlyGlnSerGlnLysValGlnGlnAlaSer 20
Db      866 ATGAATGCTGTGGAAGAGAACCAAGCCCTTGTGAGAGTCTCAAGAGGTGAGAGAGCCACG 925
Qy      21 ProProAlaValAlaGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      926 CCTCGTGTGTGACAGAGCCAGGACCTCTTCTCCAGTGTGGCCACACCCCTTGAG 985
Qy      41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
Db      986 CCAAGTAGAGGAGGATGCTGGGAGCAAGATGCTACCAAGCCGACGATGAGCTTGAG 1045
Qy      61 TyrGluAspGlyArgGlyPheGlyIleGlyGlnLysValTyrGlyLysLeuArgGlyPhe 80
Db      1046 TATGAGATGGGCGGGGCTTGTGCAATTTGAGAGCTGTGTGGGGAACCTTGGGGCTTC 1105
Qy      81 SerTyrTyrProGlyArgIleValSerTyrTyrMetThrGlyArgSerArgAlaGln 100
Db      1106 TCCTGTGTGCGAGGCGCAATGTGTCTGTGATGACAGCGCGAGCCGAGACGCTGAA 1165
Qy      101 GlyThrArgTyrValMetTyrPheGlyAspGlyLysPheSerValValCysValGlyLys 120
Db      1166 GGCATCTGCTGTGATGATGTGTTCGAGATGGCAAGTCTCAGTGTGTGTGGGAAG 1225
Qy      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db      1226 CTCATGCGCGCTGAGCTCTTCTGCGATGCTTCCACGAGCCACCTCAACAAGCAGCCCC 1285
Qy      141 MetTyrArgLysAlaIleTyrGlnValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      1286 ATGTACCGGAAGCCATCTACGAAGTCTCCAGAGTGGCAGAGGCGGTCCGGGAAGCTG 1345
Qy      161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGlnLysAsn 180
Db      1346 TTTCAGCTTGCATGACAGTATGAAGTGAACAGTGGCAAGCTGTGTGAAGTGCAGAAC 1405

```

QY 181 LysGlnMetIleGlnTrrPalaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGln 200
 DB 1406 AACGACATGATTAATGAGCCCTCGGTGCTCCAGCCCTCGGATCTTAAGAGCCCTGGAG 1465
 QY 201 ProProGlnGlnGlnLysAsnProTyrLysGlnValTyrThrAspMetTrpValGlnPro 220
 DB 1466 CCACCAAGAGAGAGAGAAATCCTTACAGAGAAAGTTTACACCAATGATGGTGGAGCCT 1525
 QY 221 GlnAlaAlaIleTrrAlaProProProProAlaLysLysProArgLysSerThrThrGln 240
 DB 1526 GAAGCAGCTCTTAACGCCCAACCCCAAGCAAGAAACCAAGAAACCAAGCAAGAG 1585
 QY 241 LysProLysValLysGlnIleLeaSPGluArgThrArgGluArgLeuValTyrGlnVal 260
 DB 1586 AAACCTAAGGTCAAGGATCATTAATGATGAGCCCAAGGAGCGGCTGATGAGGTG 1645
 QY 261 ArgGlnLysSerArgAsnIleGlnAspIleSerCysGlySerLeuAsnValThr 280
 DB 1646 CGCCAGAGTGCAGAAACATCGAGACATTTGATCTCATGTGGAGACCTCAATGTCACC 1705
 QY 281 LeuGlnHISProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGln 300
 DB 1706 CTGGACACCACTCTTCACTTGAAGGATGTCAGAGACTGAAGAACTGCTTGGAG 1765
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 DB 1766 TGTGCTTACAGATATGACGAGATGGGTACAGCTCATTCATTCGCTGCTGAGG 1825
 QY 321 ArgGlnValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGlnCysVal 340
 DB 1826 CGTGAAGTGTCAATGTGTGGGAAACAACTGCTGCGAGGTGTTGTGTGAGTGTG 1885
 QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGlnAspProTrrPasnCys 360
 DB 1886 GATCTCTGTGGGGCCAGAGAGCTGCTCAAGCAAGCATTAAAGAAAGCCCTGGAAGTGC 1945
 QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGlnAspTrpProSer 380
 DB 1946 TACATGTGCGGGCATTAAGGCGACCTTAATGGCTGTGGAAGAGCGGGAGCTGGCTTCT 2005
 QY 381 ArgLeuGlnMerPhePheAlaAsnAsnHisAspGlnGlnLysPheAspProLysValTyr 400
 DB 2006 CGACTCCAGATGTTCTTGGCAATTAACCAATGACCAAGAAATTTGACCCCAAGGTTTAC 2065
 QY 401 ProProValProAlaGlnLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2066 CCACTGTGCTGACCTGGAAGAGAGAACCCATCCGGCTGCTCTTGTGATGGAT 2125
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 DB 2126 GCTACAGGGCTCTGTGCTGGAAGGACCTGGCATCCCAAGTGAAGCCCTTCACTTGGCTCC 2185
 QY 441 GlnValCysGlnLysPheSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2186 GAGGTGTGATGAGATCTCATGACGTGGCATGTGGGCAACAGGAAAGATCATGTAC 2245
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGlnLysTrrGlyProPheAspLeu 480
 DB 2246 GTGGGGGACCTCCGAGCGTCACACAGACATATCCAGAGGTGGGCCCATTCGACCTG 2305
 QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2306 GTGATTGGAGGACATCCCTGCAATGACCTCTCCATTTTCAACCTGCGCCCAAGGACCT 2365
 QY 501 TyrGlnGlyThrGlyArgLeuPhePheGlnPheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2366 TATGAGGGTACTGCGCCCTCTTCTTGAAGTTCTACCGCTCTCGCATGATGCGGCGCC 2425
 QY 521 LysGlnGlyAspAspArgProPhePheTrpLeuPheGlnLysValValAlaMetGlyVal 540
 DB 2426 AAGGAGGAGATGATCCCTTCTGTGCTCTTGAAGATGTGATGCGCATGCGGCTT 2485

QY 541 SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys 560
 DB 2486 AGTGACAAAGGAGACATTCGCGATTTCTTGAGTCTTAACCCGCTGATGATTAACCCCAA 2545
 QY 561 GlnValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
 DB 2546 GAAAGTCTGCTGCACACAGGCGCGGTACTTCTGGGGTAACTTCTCGCATGAACAG 2605
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnLeuGlnGlyCysLeuGlnHisGlyArg 600
 DB 2606 CTTTGGCATTCACCTGTGATGATTAAGCTGAGCTGCAAGATGTCTTGAGAGACGGCAGA 2665
 QY 601 IleAlaLysPheSerLysValArgThrIleThrTrpArgSerAsnSerIleLysGlnGly 620
 DB 2666 ATAGCCAACTTACGAAAGTGAAGACATTACCAACAGGTCAAACTTTAAAGCAGGCG 2725
 QY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTrpCysThrGln 640
 DB 2726 AAAGACACGATTTCCCGCTTTCATGAACGAAAGAGACATCCTGTGTGACACTGAA 2785
 QY 641 MetGlnArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2786 ATGAAAGGGTGTGTTGGCTTCCCGTCACTACAGAGAGTCTTCAACATGAGCGCTTG 2845
 QY 661 AlaArgGlnArgLeuGlnLysArgSerTrpSerValProValIleArgHisLeuPheAla 680
 DB 2846 GGAAGGCAAGACTGTGGGCGGATCGTGAGACGTCGCGGTATCCGCCACCTTCCGCT 2905
 QY 681 ProLeuLysGlnTyrPheAlaCysVal 689
 DB 2906 CCGTGAAGAAATATTTCTTGTGTG 2932

RESULT 5 US-10-623-813-84

; Sequence 84, Application US/10623813
 ; Publication No. US2004023497A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, En
 ; APPLICANT: Okano, Masaki
 ; APPLICANT: Xie, Shaoping
 ; APPLICANT: Chen, Taiping
 ; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Use
 ; TITLE OF INVENTION: Thereof
 ; FILE REFERENCE: 0609,4560003
 ; CURRENT APPLICATION NUMBER: US/10/623,813
 ; CURRENT FILING DATE: 2003-07-22
 ; PRIOR APPLICATION NUMBER: US 09/720,086
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: PCT/US99/14373
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US 60/090,906
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: US 60/093,993
 ; PRIOR FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 119
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 84
 ; LENGTH: 2371
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-623-813-84

Alignment Scores:

Pred. No.: 0 Length: 2371
 Score: 3714.00 Matches: 679
 Percent Similarity: 98.8% Conservative: 2
 Best Local Similarity: 98.5% Mismatches: 8
 Query Match: 98.8% Indels: 0
 DB: 9 Gaps: 0

US-10-623-813-85 (1-689) x US-10-623-813-84 (1-2371)

QY 1 MetAsnAlaValGlnGlnLysGlnLysArgGlyGlnSerGlnLysValGlnGlnAlaSer 20


```

Db      217 ATAAATGCTGTGGAAGAAAACAGAGGCCCCGGGAGCTCTCAAGAGTGGAGAGGCCAGC 276
Qy      21  ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      277 CCTCGTCTGTGAGCAGGCCACCTGACCCCGCATCCCTCACTGTGGTCAACGACCTGAG 336
Qy      41  ProValGlyValAspAlaGlyAspValAsnAlaThrTyrValAlaAlaAspAspGluProGlu 60
Db      337 CCGGTGGGGTCCGATGCTGGGGACAGAAATGCCAAGACAGCCGATGCGAGCCAGAG 396
Qy      61  TyrGluAspGlyValArgGlyPheGlyIleGlyValLeuValTyrGlyValLeuArgGlyPhe 80
Db      397 TACGAGACGCGCGCGGCTTGGCATTTGGGAGCTGGTGTGGGGAAACCTGGGGCTTC 456
Qy      81  SerTyrTrpProGlyValArgIleValSerTyrTrpMetThrGlyValArgSerAlaAlaGlu 100
Db      457 TCTGTGGTGGCCAGAGCGCATTTGTCTGTGGATGACGGGCGGAGCCGACGACTGAA 516
Qy      101 GlyThrArgTyrValMetTyrPheGlyValAspGlyValPheSerValValCysValGlyVal 120
Db      517 GGCACCCGCTGGGTCAATGTGTTCCGAGACGCAAAATTCACAGTGTGTGTGAGAAAG 576
Qy      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db      577 CTGATCCCGCTGAGCTCGTTTTCAGTGCCTTCCACAGGCCACGTAACAAGCAGCC 636
Qy      141 MetTyrArgValAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      637 ATGTACCGCAAGACCATCTACGAGGTCCTGACAGTGGCCAGCGCGCGGGAAAGCTG 696
Qy      161 PheProAlaCysHisAspSerAspGluSerAspSerGlyValAlaValGluValGlnAsn 180
Db      697 TTCCTGGTGTGCACACAGCATGAGTACGACACGCCAGGCGCTGGAGGTGCAGAAC 756
Qy      181 LysGlnMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      757 AAGCCCATGATTAATGAGGCCCTCGGGGGCTTCACCTTCCTGGCCCTAAGGCCCTGAG 816
Qy      201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTyrValGluPro 220
Db      817 CCACCGAAGAGAGAGAAATCCCTCAAGAGTATACACGACATGTGTGGTGAACCT 876
Qy      221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
Db      877 GAGGACGCTGCTACCGCACCTCCACAGCCAAAGCCCGGAGAGACAGCGGAG 936
Qy      241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db      937 AAGCCCAAGGTCAAGAGATTAATGATGAGCCGACAAAGAGAGGGCTGGTGTACAGGTG 996
Qy      261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      997 CGGCAAGAGTGC CGGAAACATTGAGACATCTGCATCTCTGTGGAGGCTCAATGTTACC 1056
Qy      281 LeuGluHisProLeuPheIleGlyValMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1057 CTGGAAACACCCCTCTTGTGGAGGAATGTCCAAACCTGCAAGAACCTTTCTTGAGAG 1116
Qy      301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1117 TGTGCGTACCACTACGACACAGCAGCGGCTACCACTCTGACCACTCTGCTGTGGGGG 1176
Qy      321 ArgGluValLeuMetCysGlyValAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1177 CGTGAAGTCTCATGTGCGGAAACAAACATCTGCTGAGGTGCTTTGTGCTGAGTGTGTG 1236
Qy      341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProThrPheCys 360
Db      1237 GACCTCTTGTGTGGGGCGGGGGCTGCCAGGCGCATTAAGAAACCCCTGGAACTGC 1296
Qy      361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTyrProSer 380

```

```

Db      1297 TACATGTGGGGACACAAGGTAACCTACGGGCTGCTGGCGGGGAGAGAGACTGCCCTCC 1356
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProPolysValTyr 400
Db      1357 CGGCTCAAGATGTTCTTCCCTAATAACACAGCACAGAAATTTGACCTCCAAAGTTTAC 1416
Qy      401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      1417 CACCTGTCCCACTGAGAAAGAGAGGCCATCCGGGTCTGTCTCTTTATATGAAATC 1476
Qy      421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      1477 GCTACAGGGCTCTGTGTGTAAGACCTTGGGCATTCAGGTGACCGGTACATATGCTCTCG 1536
Qy      441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      1537 GAGGTGTGAGAGACTCCATCAACGTGGGACATGTGTCCGACACAGGGAAGATCATGTAC 1596
Qy      461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrProPheAspLeu 480
Db      1597 GTCCGGGAGCTCCGACGCTCACAGAGACATATCCAGAGTGGGCCCATTCGATCTG 1656
Qy      481 ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db      1657 GTGATTTGGGGACATCCCTGCAATGACCTTCATGTCACACCTGCTCGCAAGGGCCTC 1716
Qy      501 TyrGluGlyThrGlyValArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      1717 TACGAGGGCAGTCGGCGGCTCTTCTTGTGATTCACCGGCTCTGATATGCGCGCCC 1776
Qy      521 LysGluGlyAspAspArgProPhePheTyrLeuPheGluAsnValValAlaMetGlyVal 540
Db      1777 AAGAGAGAGATGATGCGCCCTTCTTGTGCTCTTGTGAATGTGTGGCCATGGCGCTT 1836
Qy      541 SerAspLysArgAspIleSerArgPheLeuGlnLysAsnProValMetIleAspAlaLys 560
Db      1837 AGTGACAAGAGGAGCATCTCGGATTTCTCGAGTCCAAACCTGTGATGATGATGCCTAA 1896
Qy      561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyValAsnLeuProGlyMetAsnArg 580
Db      1897 GAAGTGCAGCTGCACACAGGGCCCGCTACTTGGGGTAAACCTTCCCGTATGAACAG 1956
Qy      581 ProLeuAlaSerThrValAsnAspLysLeuGlnLysGlnLysCysLeuGluHisIleArg 600
Db      1957 CCGTTGGCATCCACTGTGATATGATTAAGCTGAGAGCTGAGAGATGCTGAGACATGGCAG 2016
Qy      601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2017 ATAGCCAAGTTCAGCAAAAGTGAAGACATTTACTACAGGTCAAACTCCATAAAGCAGGCG 2076
Qy      621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
Db      2077 AAGAGCCAGCATTTCTCTGTTCATGAATGAAGAAAGAGACATCTTAAGTGCATGAA 2136
Qy      641 MetGluAspValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2137 ATGAAAGGGATTTTGTGTTTCCCACTCACTTACTGACGTCTCAACACTGAGCCCTTG 2196
Qy      661 AlaArgGlnArgLeuLeuGlyValArgSerTyrSerValProValIleArgHisLeuPheAla 680
Db      2197 GCGAGGACAGAGACTGCTGGGCGGTCAATGAGCGTGCAGTCACTCGCACCTTTCGCT 2256
Qy      681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2257 CCGCTGAAGAGATATTTGCGGTGTG 2283

```

RESULT 6
 US-10-144-577-2
 ; Sequence 2, Application US/10144577
 ; Publication No. US20030083292A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MacLeod, Alan Robert
 ; TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms

FILE REFERENCE: MET-005
 CURRENT APPLICATION NUMBER: US/10/144,577
 CURRENT FILING DATE: 2002-05-13
 PRIOR APPLICATION NUMBER: US 60/290,202
 PRIOR FILING DATE: 2001-05-11
 PRIOR APPLICATION NUMBER: US 60/290,212
 PRIOR FILING DATE: 2001-05-11
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 3005
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-144-577-2

Alignment Scores:

Pred. No.:	0	Length:	3005
Score:	3714.00	Matches:	679
Percent Similarity:	98.8%	Conservative:	2
Best Local Similarity:	98.5%	Mismatches:	8
Query Match:	98.8%	Indels:	0
DB:	6	Gaps:	0

US-10-623-813-85 (1-689) x US-10-144-577-2 (1-3005)

QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnLysValGluGluAlaSer 20
 Db ATGAAATGCTGTGCAAGAAACCAAGGGCCCGGGAGCTCTCAGAGGTGGAGGGCCAGC 958
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 Db CCGCTCGCTGTGAGCAGCCACCTGACCCCGCATCCCGCTGCTACCGAGCTTGAG 1018
 QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
 Db CCGGTGGGGTCCGATGCTGGGAGCAAGATGCCCAAGCAGGCGATGACGAGCCAGAG 1078
 QY 61 TyrGluAspGlyArgGlyPheGlyTleGlyGluLeuValTrrpGlyLysLeuArgGlyPhe 80
 Db TACGAGAGCCGGCCGGCTTTGGCATTTGGGAGCTGTGTGGGAAACTGGGGGCTTC 1138
 QY 81 SerTrrpProGlyArgGlyLeuAlaSerTrrpMetThrGlyArgSerArgAlaAlaGlu 100
 Db TCTGTGGGCGAGCCGCGCATTTGTCTTGTGATGACGGGCCGAGACCGAGCTGAA 1198
 QY 101 GlyThrArgTrrpValMetTrrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 Db GGCACCCGGTGGGTGATGTGTGCGAGACGGCAAAATTTCTCAGTGTGTGTGAGAG 1258
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 Db CTATATGCCGTGAGCTGTTTTCAGTGTCTCCACGAGCCACGTCACCAACAGCACCCC 1318
 QY 141 MetTyrArgLysAlaTleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 Db ATGTATCCGCAAGCCATCTACAGAGTCTCTCAGGTGGCCAGACCCCGCGGGAACTG 1378
 QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180
 Db TTCCCGGTGTGCCACGACAGATGATGACACTCCCAAGGCGGTGAGGTGCAGAAC 1438
 QY 181 LysGlnMetTleGluTrrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 Db AACCCCTGATGATGATGGCCCTGGGGGCTTCCAGGCTTCTGGCCCTTAAGGGCTGAG 1498
 QY 201 ProProGluGluGluLysAsnProTrrpLysGluValTrrpThrAspMetTrrpValGlnPro 220
 Db CCACCAAGAGAGAGAAATCTCTACAAAGATGTCACGAGCATCTGGGTGAACT 1558
 QY 221 GluAlaAlaAlaTrrpAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
 Db GAGGACACTGCTACGACCACTCCACAGCCAAAGAGCCCGCGAAGAGACGACGCGAG 1618
 QY 1559 GAGGACACTGCTACGACCACTCCACAGCCAAAGAGCCCGCGAAGAGACGACGCGAG 1618

QY 241 LysProLysValLysGluTleTleAspGluArgThrArgGluArgLeuValTrrpGluVal 260
 Db AACCCCAAGGTCAAGAGATTTATGATGAGCGCACAGAGAGCGCTGTGTACGAGTG 1678
 QY 261 ArgGlnLysCysArgAsnTleGluAlaPheCysTleSerCysGlySerLeuAsnValThr 280
 Db CGGCAAGATGCCGAACATTGAGCATCTGATCTCTGTGGAGCTCAATGTTACC 1738
 QY 281 LeuGlnHisProLeuPheTleGlyTyrMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 Db CTGGAACACCCCTCTCTCTGTGAGGAATGTGCCAAATCTGCAGAACTGCTTCGAG 1798
 QY 301 CysAlaTrrpGluTrrpAspAspGlyTrrpGlnSerTrrpCysThrTleCysCysGly 320
 Db TGTGCTACCAAGTACAGACAGACGAGCTACAGCTCTTACCACTTGTCTGTGGGGC 1858
 QY 321 ArgGluValLeuMetCysGlyLysAsnAsnCysCysArgCysPheCysValGluCysVal 340
 Db CGTGAAGTCTCATATGCGGAAACAAACAACTGTCAGAGTCTTTTGCCTGAGTGTG 1918
 QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaTleLysGluAspProTrrpAsnCys 360
 Db GACCTTGTGTGGGCGGGGGCTGCGCCAGGCGCATTAAGGAACCCCTGGAACCTGC 1978
 QY 361 TyrMetCysGlyHisLysGlyThrTrrpGlyLysLeuValArgArgGluAspTrrpSer 380
 Db TACATGTGGGGACAAAGGTACTACGGGCTCTCGCGGGAGAGACCTGGCTTC 2038
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProLysValTyr 400
 Db CGGCTCAGATGCTTCTCGTAAATTAACAGACCAAGAAATTTGACCTCCAAAGTTTAC 2098
 QY 401 ProProValProAlaGluLysArgLysProTleArgValLeuSerLeuPheAspGlyTle 420
 Db CCACTGTCTCCACTAG 2158
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyTleGlnValAspArgTrrpTleAlaSer 440
 Db GCTACAGGGCTCTGTGTCTGAAGACTTGGGATTCAGGTGACCGCTCATTTGCTCG 2218
 QY 441 GluValCysGluAspSerTleThrValGlyMetValArgHisGlnGlyLysTleMetTyr 460
 Db GAGGTGTGAGGACTCCATCACGAGTGGCATGTGCGGACCAAGGAGAGATCATCTAC 2278
 QY 461 ValGlyAspValArgSerValThrGlnLysHisTleGlnGluTrrpGlyProPheAspLeu 480
 Db GTGGGAGAGCTCCGACGCTCACAGAGACATATCCAGAGTGGGGCCCATTCGATCTG 2338
 QY 481 ValTleGlySerProCysAsnAspLeuSerTleValAsnProAlaArgLysGlyLeu 500
 Db GTATTTGGGGGAGTCTCTGCAATGACCTTCATGTCTCAACCTCTGTCGAAGGGCTC 2398
 QY 501 TrrpGluGlyThrGlyArgLeuPhePheGluPheTrrpArgLeuLeuHisAspAlaArgPro 520
 Db TACGAGGAGACTGGCGGCTCTCTTGAATTTCACTGCTCTGATGATGCGGCGCC 2458
 QY 521 LysGlnGlyAspAspArgProPhePheTrrpLeuPheGluAsnValAlaMetGlyVal 540
 Db AAGGAGGAGATATGCCCTTCTCTGCTTTGAGATGTGGGCGCAAGGGGCTT 2518
 QY 541 SerAspLysArgAspTleSerArgPheLeuGluLysSerAspProValMetTleAspAlaLys 560
 Db AGTGACAAAGAGGACATCTGCGATTTCTGAGTCCAAACCTGTGATGATGACCAA 2578
 QY 561 GluValSerAlaAlaHisArgAlaArgTrrpPheTrrpGlyAsnLeuProGlyMetAsnArg 580
 Db GAAGTGTACGCTCAACAGGCGCCGCTACTTCTGGGGTAACTTCCCGGTAAAGAACG 2638
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGluCysLeuGluHisGlyArg 600
 Db CCGTTGCAATCCACTGTGATGATTAAGCTGAGACTCAAGAGTGTCTGGAGCATGGCAG 2698
 QY 601 TleAlaLysPheSerLysValArgThrTleThrThrArgSerAsnSerTleLysGlnGly 620

```

Db      2699 ATAGCCAACTTCACGAAATGAGACCACTTACTACAGAGTCAAACTCCATTAAGCGGGC 2758
Qy      621 LysaSPG1NHISpHeProValPheMeAaNGluLysGluasp1LeuTPrCyStrGlu 640
Db      2759 AAAGACACAGCATTTCTCTCTCTTCATGAATGAGAAAGAGACATCTTATGTCACATGAA 2818
Qy      661 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAspMetSerArgLeu 660
Db      2819 ATGGAAAGGGATTTGGTTTCCAGTCCACTACTAGACGTCTCCAAACATGAGCGCTTG 2878
Qy      661 A1aAGG1nArgLeuLeuGlyArgSerTPSerValProValI1earGHisLeuPheAla 680
Db      2879 GCGAGGACAGACCTGCGGCGCGGTATGAGACGTCACAGTCACTCCGACCTCTTCGCT 2938
Qy      681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2939 CCGCTGAAGAGATATTGTCGTGTG 2965

```

RESULT 7

```

US-10-172-118-132
; Sequence 132, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 132
; LENGTH: 3005
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF067972
; DATABASE ENTRY DATE: 2001-06-18
; US-10-172-118-132

```

Alignment Scores:

```

Pred. No.: 0 Length: 3005
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
DB: 7 Gaps: 0

```

US-10-623-813-85 (1-689) x US-10-172-118-132 (1-3005)

```

Qy      1 MetaEna1aValGluGluAsnGln1aSerGlyGluSerGlnLysValGluGlu1aSer 20
Db      899 ATGAATGCTGTGGAAGAAAACCAAGGGGCCCGGAGATCTCAGAAAGTGAGAGAGCCAGC 958
Qy      21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      959 CCTCTGCTGTGCACAGCCCACTGACCCCGCATCCCCACAGTGTGCTAACCAAGCCTGAG 1018
Qy      41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
Db      1019 CCGGTGGGCTCCGATCTGTCGGGACAAAGATGCCCAAGACAGCGCATGACAGCCAGAG 1078
Qy      61 TyrGluAspGlyArgGlyPheGlyI1eGlyGluLeuValTPrGlyLysLeuArgGlyPhe 80
Db      1079 TACGAGACGCGCGGGCTTGTGGCATTTGGGAGAGTGTGTGGGGAAGAACTCGCGGGCTTC 1138

```

```

Qy      81 SerTrpTrpProGlyArgI1eValSerTrpTrpMetThrGlyArgSerArgAlaI1aGlu 100
Db      1139 TCCTGTGACCAAGCGCGCATTTGTGTGTGTGAGAGAGCGGACCGGACCACTGA 1198
Qy      101 GlyThrArgTrpValMetTrpPheGlyLysPheSerValValCysValGluLys 120
Db      1199 GGCACCCGCTGGGTGATGTGGTGGAGACCGCAAAATCTCAGTGTGTGTGTGAGAG 1258
Qy      121 LeuMetProLeuSerSerPheCysSerAlaPheI1eGlnAlaThrTyrAsnLysGlnPro 140
Db      1259 CTGAGCCGCTGAGCTCTCTTTTTCAGTCCCTCCACAGGCCACCTACAAAGAGAGCCC 1318
Qy      141 MetTyrArgLysAlaI1eTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      1319 ATGATCCGAAAGCCCATCTACAGAGTCTCGAGTGGCCAGCAGCCCGCGGGAAGCTG 1378
Qy      161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluGlnAsn 180
Db      1379 TTCCCGGTGTGCACAGACAGCGATGAGAGTGACATGCCAAGCGCGTGGAGGTGCAGAG 1438
Qy      181 LysGlnMetI1eGluTPrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      1439 AAGCCATGATTGAATGGGCCCTGGGGGCTTCCAGCCTTCTGGGCCCTAAAGGCCCTGAG 1498
Qy      201 ProProGluGluGluLysAsnProThrLysGluValTyrThrAspMetTrpValGluPro 220
Db      1499 CCACCAAGAAAGAAAGAAATCCCTACAAAGAGTGCACCGCATGTGGGTGAACTT 1558
Qy      221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrglu 240
Db      1559 GAGGAGGTGCTTACGCAACCACTCCACAGCCAAAGAGCCCGGAAAGACAGCGAG 1618
Qy      241 LysProLysValLysGluI1eLeuAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db      1619 AAGCCCAAGGTCAAGAGATATTATGATGAGCGCACAAAGAGCGGCTGTGTACAGAGTG 1678
Qy      261 ArgGlnLysCysArgAsnI1eGluAspI1eCysI1eSerCysGlySerLeuAsnValThr 280
Db      1679 CGGCAGAGTGCAGGACATTTGAGACATCTGCACTCTCTGTGAGAGCTCAATTGACC 1738
Qy      281 LeuGluHisProLeuPheI1eGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1739 CTGAAACACCCCTCTTCTGTGAGAGATGTGCCAAAACCTGAAGACGCTTCTCGAG 1798
Qy      301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrI1eCysGlyGly 320
Db      1799 TGTGCTGTACAGTACGACGACAGCGCTTACCACTCTGACCATCTGCTGTGGGGC 1858
Qy      321 ArgGluValLeuMetCysGlyLysAsnAsnAsnCysArgCysPheCysValGluCysVal 340
Db      1859 CGTGAAGTGTCTCATGTGTGGGAAACAACAACCTGTCTCAGAGTCTTTTCCGTGAGGTGTG 1918
Qy      341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaI1eLysGluAspProTrpAsnCys 360
Db      1919 GACCTCTTGTGTGGGCGCGGGGCTGCCAGGAGCCATTAAAGGAAGACCCCTGGAATCC 1978
Qy      361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpSer 380
Db      1979 TACATGTGCGGGACCAAGGATACCTACCGGCTGTGCGCGCGGAGAGATCGGCCCTCC 2038
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db      2039 CGGCTCCAGATGTTCTTCTGCTTAACCAACACACAGAAATTTGACCTCCAAAGTTTAC 2098
Qy      401 ProProValProAlaGluLysArgLysProI1eArgValLeuSerLeuPheAspGlyI1e 420
Db      2099 CCACCTGTCCAGCTGAAGAAAGAGAGCCCATCCGGGTGTGTCTCTCTTATGAAATTC 2158
Qy      421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyI1eGlnValAlaArgTyrTrpLeuAsn 440
Db      2159 GCTACAGGGCTTCGTGTGCTGAAGAGATTGGGCAATTCAGGTGAGACCGCTACATTCCTCG 2218

```

QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 2219 GAGGTGTGTAGAGATCTCATACAGTGGGATGTGGGACCAAGGGAAGATCATGTAC 2278
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGlnTyrPheProPheAspLeu 480
DB 2279 GTGGGGGACGTCCGACGCGTCAACAGAACATATCCAGAGAGTGGGGCCATTGCATCTG 2338
QY 481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 2339 GTGATTGGGGGACGTCCCTCATGTACCTCTCCATGCTCAACCTGCTCCCAAGGCGCTC 2398
QY 501 TyrGlnGlyThrCysArgLeuPhePheGlnPheThrArgLeuLeuHisAspAlaArgPro 520
DB 2399 TACGAGGGCATGCGCGGCTCTTGTGAGTTTACCGCTCTCGCATGTATGCGGCGCC 2458
QY 521 LysGlyGlyAspAspArgProPhePheTrpLeuPheGlnAsnValAlaAlaMetGlyVal 540
DB 2459 AAGGAGGGAGATATCCGCCCTTCTTGCGCTCTTGAGAAATGGTGGCCATGGGCGTT 2518
QY 541 SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys 560
DB 2519 AGTGACAAGAGGACATCTCGCATTTCTCGAGTCCCAACCTGTGATGATGACCAAA 2578
QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
DB 2579 GAAGTGTCAAGTGCACACAGGCGCGCTACTTCTGGGGTAACTTCCGGTATGAACAG 2638
QY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnGlnGlnLysLeuGlnHisIleGlyArg 600
DB 2639 CCGTTGGCATCCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2698
QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
DB 2699 ATAGCCCAAGTTCAGCAAGTGAAGGACCATTAATCAAGGTCCTCAATCAAGGCGG 2758
QY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTrpCysThrGln 640
DB 2759 AAGGACGAGATTTCTGCTTCAATGAAAGAAAGAGACATCTTATGGTCACTGAA 2818
QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
DB 2819 ATGGAAAGGATATTTGGTTTCCAGTCCACTATACGACGTCCCAACATGAGCCCTTG 2878
QY 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
DB 2879 GCGAGGCGAGGACTGCTGGGCGCGTATGAGCGGTGCAGTCCGACCTTCTTGCT 2938
QY 681 ProLeuLysGlnTyrPheAlaCysVal 689
DB 2939 CCGCTGAAGGAGTATTTGCGGTGTG 2965

RESULT 8
US-10-342-887-132
Sequence 132, Application US/10342887
Publication No. US20040058340A1
GENERAL INFORMATION:
APPLICANT: Dai, Hongyue
APPLICANT: He, Yundong
APPLICANT: Linsley, Peter S.
APPLICANT: Mao, Mao
APPLICANT: Roberts, Christopher J.
APPLICANT: Van de Veer, Laura Johanna
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-188-999
CURRENT APPLICATION NUMBER: US/10/342,887
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/298,918
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/380,710
PRIOR FILING DATE: 2002-05-14

PRIOR APPLICATION NUMBER: 10/172,118
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 132
LENGTH: 3005
TYPE: DNA
ORGANISM: Homo sapiens
US-10-342-887-132

Alignment Scores:
Pred. No.: 0 Length: 3005
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
Gaps: 0

US-10-623-813-85 (1-689) x US-10-342-887-132 (1-3005)

QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGlnSerGlnLysValGluGluAlaSer 20
DB 899 ATGAATGCTGTGAGAAAGAAACCAAGGGGCCCGGGAGTCTCAAGATGAGAGGCGCAC 958
QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 959 CCTCTGCTGTGCACACAGCCCACTGACCCCGCATCCCACTGCTGCTACACGCTGAG 1018
QY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
DB 1019 CCGGTGGGTCCGAGTGTGGGAGCAAGATGCCCAAGACAGGCGATGACGAGCGCAGAG 1078
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
DB 1079 TACGAGAGGCGCGGGGCTTGGCATGTGGAGCTGTGTGGGAAACTGCGGGCTTC 1138
QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaIleGlu 100
DB 1139 TCCTGTGGCCAGGCGCGCATTTGTCTGTGGATGACGCGCGGACGACACTGAA 1198
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValAlaCysValGlnLys 120
DB 1199 GGCACCGCGTGGTATGTGTTCGAGACGCGCAATTTCTAGTGTGTGTAGAG 1258
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 1259 CTGATGCGCTGAGCTGTTTTCAGTGTCTTCCACAGGCCATCAACAAGACGCC 1318
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 1319 ATGTACCGCAAGGACATCTAGAGTCTGACAGTGGCGCAGAGCGCGGGGAAGCTG 1378
QY 161 PheProAlaCysHisAspSerAspGlnSerAspSerGlyLysAlaValGluValAsn 180
DB 1379 TTCCCGGTGTGCACAGCAGCATAGAGTGCACCTGCCAAGGCGGTGAGGTGAGAAC 1438
QY 181 LysGlnMetIleGlnTrpAlaLeuGlyPheGlnProSerGlyProLysGlyLeuGln 200
DB 1439 AAGCCCATGATTGATGGCCCTTGGGGGCTTCCAGCTTCTGCGCCTTAAGGCGCTGAG 1498
QY 201 ProProGluGluGlnLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
DB 1499 CCACCAAGAAAGAAAGATCCCTACAAAGAGTACACGACATGTGGGTGAAACT 1558
QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
DB 1559 GAGGAGGTGCTTACGACCACTTCCACAGCAAAAGCCCGGAAAGGACACCGGAG 1618
QY 241 LysProLysValLysGlnIleIleAspGluArgTyrArgGluArgLysValTyrGluVal 260
DB 1619 AAGCCCAAGTCAAGAGATTTTATGAGCCCAAGAGAGCGGCTGTGTACAGAGTG 1678
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280

```

Db      1679 CGGCAAGAGTCGGCAACATTGAGCATCTCGTGTGGAGCCCTCAATGTTACC 1738
Qy      281 LeuGluHisProLeuPheIleGlyGlyMetCysGlnAsnCysIysAsnCysPheLeuGlu 300
Db      1739 CTGGAAACACCCCTCTGTTGGAGAAATGTCMAAACCTGCAAGAACCTGTTTCGGAG 1798
Qy      301 CysAlaTyrGlnTyrAspAspArgIleTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1799 TGTGCTACCACTAGACACACACGCGCTACCACTCTGACCACTCTGCTGTGGGGG 1858
Qy      321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGlyCysVal 340
Db      1859 CGTGAAGTCTCACTGTGGCAACCAACTGTCAGAGGCTTTGGCGTGGAGTGTG 1918
Qy      341 AspleuLeuValGlyProGlyAlaIleGlnAlaIleIleGlyGlnAspProTrpAsnCys 360
Db      1919 GACCTCTTGTGGGGGGGGGGGGCTGCCAGGCGCATTTAAGMAAACCTCGGAACCTGC 1978
Qy      361 TyrMetCysGlyHisIleGlyGlyThrTyrGlyLeuLeuArgArgGlyLeuAspTrpProSer 380
Db      1979 TACATGTGGGGGCAAGGCTACTACGGGCTCTCGGCGGGAGAGAGAGAGCCCTCC 2038
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProIleValTyr 400
Db      2039 CGGCTCCAGATGTTCTTCGCTAATAACACGACAGAAATTGACCTCCCAAGGTTTAC 2098
Qy      401 ProProValProAlaGlySarGlyProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2099 CCACCTGTCCTCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2158
Qy      421 AlaThrGlyLeuLeuValLeuLeuAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2159 GCTACAGGGCTCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2218
Qy      441 GlnValCysGlyAspSerIleThrValGlyMetValArgHisGlnGlyIleGlyIleMetTyr 460
Db      2219 GAGGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2278
Qy      461 ValGlyAspValArgSerValThrGlnIleHisIleGlnGluTyrProIlePheAspLeu 480
Db      2279 GTGGGGAGAGCTCCGAGCTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2338
Qy      481 ValIleGlyIleSerProCysAsnAspLeuSerIleValAspProAlaArgIleGlyLeu 500
Db      2339 GTGATTTGGGGGAGAGTCTCTGCAATGACCTCTCATGCTCAACCTCTCTGCAAGAGCCCTC 2398
Qy      501 TyrGlnGlyThrGlyArgLeuPheGluPheGlyArgLeuLeuHisAspAlaArgPro 520
Db      2399 TACGAGGGGACAGTGGCGGCTCTTCTTGAGTTCTACCGCTCTCATGATGGCGGCC 2458
Qy      521 LysGlnGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
Db      2459 AAGGAGGGAGATGATGCGCCCTCTTCTGAGTGTGAGAAATGTGTGGCCAGAGGGCGT 2518
Qy      541 SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys 560
Db      2519 AGTGACAAAGAGGACATCTCGCATTTCTCGAGTCAACCTCTGTATGATGACCAAA 2578
Qy      561 GlnValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      2579 GAAGTGTCACTGACACAGAGGGCCGCTACTTGTGGGTAACTTCCCGGTATGACAGAG 2638
Qy      581 ProLeuAlaSerThrValAsnAspLysLeuGlnGlnGlyCysLeuGlnHisGlyArg 600
Db      2639 CCGTGGGCACTCACTGTGATGATTAAGCTGAGAGCTCAGAGAGTGTGAGAGCTGGCAGG 2698
Qy      601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2699 ATAGCCAACTTACGCAAAAGTAGAGACATTAACAGAGTCAAACTCCATTAAGCAGGGC 2758
Qy      621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTrpCysThrGln 640
Db      2759 AAAGACCAACATTTCTGTCTTCACTGATGAGAAAGAGACATCTTATGTGTCACTGAA 2818

```

```

Qy      641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2819 ATGAAAGGGATTTGGTTTCCCACTCACTACTGAGCTCTCAACATGAGCCCTTG 2878
Qy      661 AlaArgGlnArgLeuGlnValArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db      2879 GCGAGGAGAGAGCTCTGGGGCGGTCAATGAGAGCGTGCAGTATCGCCACCTCTTCGCT 2938
Qy      681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2939 CCGTGAAGAGACTATTTCGCTGTG 2965

RESULT 9
US-10-330-773-279
; Sequence 279, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 279
; LENGTH: 3009
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-279

Alignment Scores:
Pred. No.: 0
Score: 3714.00
Percent Similarity: 98.8%
Best Local Similarity: 98.5%
Query Match: 98.8%
DB: 11
Length: 3009
Matches: 679
Conservative: 2
Mismatch: 8
Indels: 0
Gaps: 0

US-10-623-813-85 (1-689) x US-10-330-773-279 (1-3009)
Qy      1 MetAsnAlaValAlaGlnGlnAsnGlnAlaSerGlyGlnLysValAlaGlnGlnAlaSer 20
Db      903 ATGATGCTCTGTGAGAGAAACACAGGGGCGCGGGAGTCAAGAGTGAAGAGGCGCAC 962
Qy      21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      963 CCTCTGCTGTGACACAGCCCACTGACCCCGCATCCCACTGTGCTACCAAGCCTTGAG 1022
Qy      41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysValAlaAspAspGluProGlu 60
Db      1023 CCGTGGGGTCCGAGTGTGGGAGCAAGATGCAACCAAGCAGCGATACAGGCGCAGAG 1082
Qy      61 TyrGluAspGlyArgGlyPheGlyIleGlyGlnLeuValTrpGlyLysLeuArgGlyPhe 80
Db      1083 TACGAGAGAGCGCGGGGCTTGGCATTTGGGAGAGCTGTGTGGGGGAAACTGCGGGCTTC 1142
Qy      81 SerTrpTrpProGlyArgIleValSerTrpMetThrGlyArgSerArgAlaIleGlu 100
Db      1143 TCTGTGTGCGCAAGCGCGCATGTGTCTGTGTGATGACGCGCGGAGCGGACAGCTGAA 1202
Qy      101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGlnLys 120
Db      1203 GGCACCCCGTGGGTATGTTGTTGAGAGCGGCAATTTCTGCTGTGTGTGTTGAAAG 1262
Qy      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db      1263 CTGAGCGCGCTAGCTCGTTTGGAGTGTGCTCCACAGGCGCACGTAACAAGCAGGCC 1322
Qy      141 MetTyrArgValAlaIleTyrGlnValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      1323 ATGTACCGCAAGGCACTACAGAGTCTGTGAGGTGGCCAGCAGCGCGCGGGGAGAGCTG 1382

```

```

QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyValAlaGluValGlnAsn 180
Db 1383 TTCCCGGTGGCCACGACGATGAGATGACCTCCAAAGGCCGTGGAGGTGCAGAAC 1442
QY 181 LysGlnMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db 1443 AACCCCATGATTAAATGGCCCTGGGGGGCTTCAGGCTCTTGCCCTTAAGGGCCTGGAG 1502
QY 201 ProProGluGluGluLysAsnProGlyLysGlyValAlaTrpThrAspMetTrpValGluPro 220
Db 1503 CCACCAAGAAAGAAAGAAATCCCTTAAGAAAGTGAACAGGACATGTGGGTGAACCT 1562
QY 221 GluAlaAlaAlaAlaTrpAlaProProProAlaLysLysProArgLysSerThrTrpGlu 240
Db 1563 GAGGCGAGCTGCTTACGACCACTCCACGCAAAAAGCCCGGAGAGACACAGCCGAG 1622
QY 241 LysProLysValLysGlyIleIleAspGluArgThrArgGluArgLeuValAlaTrpGluVal 260
Db 1623 AACCCCAAGGTCAAGGATTTGATGAGCGCAAGAGAGAGGGCTGGTGTACAGAGTG 1682
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 1683 CGGCAGAAAGTCCGGAACATTGAGACATCTGCATCTCTGTGGAGCCTCAATGTACC 1742
QY 281 LeuGlnHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db 1743 CTGGAAACACCCCTCTCTGTGGAGAAATGTGCCAAATGCAAGAACTCTTCTGGAG 1802
QY 301 CysAlaIleTrpGlnTrpAspAspAspGlyTrpGlnSerTrpCysThrIleCysCysGly 320
Db 1803 TGGCGGTACAGTACGACGAGAGAGGCTTACCATCTCACTGACCATCTCTGTGGGGC 1862
QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1863 CGTGAGGTGCTCATGTGCGGAAACAACTGTGCGAGGTCTTTCGTGGAGTGTG 1922
QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db 1923 GACCTCTGTGGTGGCGGGGGGCTGCCAGGCAAGCCATTAAAGAGAGCCCTTGAAATGC 1982
QY 361 TyrMetCysGlyHisLysGlyTrpTrpGlyLeuLeuArgArgArgLysAspTrpProSer 380
Db 1983 TACATGTGGCGGACAAAGGTACTTACGGGCTGTGGCGCGGAGAGACTGGCCTTC 2042
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db 2043 CGGCTCCAGATGTTCTTCGCTAATTAACGACGACAGAAATTTGACCTCCAAAGTTTAC 2102
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 2103 CCACCTGTCCCACTGAGAGAGAGAACCCATCCGGGTGTCTCTCTTGTGATGATC 2162
QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTrpIleAsnSer 440
Db 2163 GCTACAGGGCTCTGTGTCTGAAAGACTTGGGCACTTACGTGACCGCTACATTGCTTCG 2222
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 2223 GAGGTGTGTAGGACTCATCGGTGGCATGTGTGGGACACGAGGGAAGATCATGTAC 2282
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db 2283 GTGGGGGACCTCGCAGCGTCAACAGAACATATCCAGAGTGGGGCCCATTCGATCTG 2342
QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaGlyLysGlyLeu 500
Db 2343 GTGATTGGGGGCACTCCCTGCAATGACTCTCATCTCAACCTGCTCCGAGGGGCTC 2402
QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTrpArgLeuLeuHisAspAlaArgPro 520
Db 2403 TACGAGGCGACTGGCGCTCTTCTTGAATTAACGCTCCGTCATGATGCGCGGCC 2462

```

```

QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
Db 2463 AAGAGGGAGATGATGCCCCCTTCTTCTGCTCTTTGAGAAATGTGTGGCCATGGCCCTT 2522
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 2523 AGTCAAGAGGAGACATCTCGGATTTCTCGATTCACCACTCTGTGATGATGCGCAA 2582
QY 561 GluValSerAlaAlaHisArgAlaArgTrpPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db 2583 GAAGTGTGACTGTCACACAGGGGCCCTTACTCTGTGGGTAACTTCCCGGTATGAACAG 2642
QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGlyCysLeuGlnHisGlyArg 600
Db 2643 CGTTGGCATCCACGTGTAATGATTAAGCTGAGACTGCAAGAGTCTTGAGACATGGCAG 2702
QY 601 IleAlaLysPheSerLysValArgThrIleThrArgSerAsnSerIleLysGlnGly 620
Db 2703 ATAGCCAACTTACAGCAATGAGAGACATTACTACAGAGTCAAACTCCATTAAGCAGGG 2762
QY 621 LysAspGlnHisPheProValPheMetLeuGluLysGluAspIleLeuTrpCysThrGlu 640
Db 2763 AAAGACCACTTTCTCTCTTCAATGAGAAAGAGACATTTATGTGCTGAA 2822
QY 641 MetGluArgValPheGlyPheProValHisTrpThrAspValSerAsnMetSerArgLeu 660
Db 2823 ATGGAAGGGATTGTGTTCCAGTCCACTTACGATGATCTTCCCAATGAGCCCTTG 2882
QY 661 AlaArgLysArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2883 GCAGGCAAGACTGTGGCGCGGTGATGAGCGTCCAGTATCCGCCACCTTCTGCT 2942
QY 681 ProLeuLysGluTrpPheAlaCysVal 689
Db 2943 CGCTGAAGAGTATTTGCTGTGTG 2969

```

RESULT 10

US-09-720-086-3

Sequence 3, Application US/09720086

Publication No. US20060084053A1

GENERAL INFORMATION:

APPLICANT: LI, En

APPLICANT: Okano, Masaki

TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,

FILE REFERENCE: 0609.456002

CURRENT APPLICATION NUMBER: US/09/720, 086

PRIOR FILING DATE: 2000-12-20

PRIOR APPLICATION NUMBER: PCT/US99/14373

PRIOR FILING DATE: 1998-06-25

PRIOR APPLICATION NUMBER: 60/093, 993

PRIOR FILING DATE: 1998-07-24

NUMBER OF SEQ ID NOS: 82

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 4293

TYPE: DNA

ORGANISM: Homo sapiens

US-09-720-086-3

Alignment Scores:

Pred. No.: 0 Length: 4293

Score: 3714.00 Matches: 679

Percent Similarity: 98.8% Conservative: 2

Best Local Similarity: 98.5% Mismatches: 8

Query Match: 98.8% Indels: 0

DB: 5 Gaps: 0

US-10-623-813-85 (1-689) x US-09-720-086-3 (1-4293)

```

OY 1 MetAsnAlaValGluGluGlnGlnAlaSerGlyGluSerGlnValGluGlnAlaSer 20
DB ATGAATGCTGTGGAAAGAAACAGAGGGCCCGGGAGCTCAGAAAGTGGAGAGGCCAGG 958
OY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB CCTCCGCTGTGGACAGAGCCACTGACCCCGCATCCCGCACTGTGGCTACAGAGCTGAG 1018
OY 41 ProValGlyGlyAspAlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGlu 60
DB CCGGTGGGGTCCGATCTGGGAGCAAGATGCCAACAGAGCCAGTCAGAGCCAGAG 1078
OY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrrpGlyLysLeuArgGlyPhe 80
DB TACGAGGAGCGCCGGGGCTTTGGCATTTGGGAGCTGTGTGGGGAACCTCGGGGCTTC 1138
OY 81 SerTrrpTrrpProGlyValArgIleValSerTrrpTrrpMetThrGlyValSerArgAlaIleGlu 100
DB TCTGTGGTGGCCAGGCGCGCATTTGTGTCTTGTGGATGACGGGCGCGAGCCGAGCAGCTGAA 1198
OY 101 GlyThrArgTrrpValMetTrrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB GGACCCCGCTGGGTGTCATGTGGTTCGGAGACGGCAAAATCTCAGTGGTGTGTGGAAG 1258
OY 121 LeuMetProLeuSerSerPheCysSerAlaPheIleGlnAlaThrTyrAsnLysGlnPro 140
DB CTGATGCGCCTGAGCTCGCTTTTGCAGTGGCTTCCACAGGCCAGTCACAAAGCAGCC 1318
OY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB ATGTACCCGAAAGCCATCTACGAGGCTCTGACAGGTGGCAGAGCGCGCGGGAAGCTG 1378
OY 161 PheProAlaCysHisAspSerAspGlyLysAspSerGlyLysAlaValAluGlnAsn 180
DB TTTCCGCTGTGTCACACAGCATGAGAGTGAACATCCAGCCAGGCGCTGAGAGTGCAGAAC 1438
OY 181 LysGlnMetIleGluTrrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB AAGCCATGATTTGAATGGGCCCTCGGGGGGCTTCCAGCCTTTCGGCCCTTAAGGGCCCTGAG 1498
OY 201 ProProGluGluGlnLysAsnProTyrLysGluValTyrThrAspMetTrrpValGluPro 220
DB CCACCAAGAGAGAGAGATCCCTACAAAGAGTGTACACGACATGTGGGTGGAACCT 1558
OY 221 GlnAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrArgLys 240
DB GAGGAGGCTGCTACGACACCTCCACAGCCAAAGCCCGGAGAGCAGCGGGAG 1618
OY 241 LysProLysValLysGlyLysIleAspGluArgThrArgGluValTyrGluVal 260
DB AAGCCCAAGTCAAGAGATTTATGATGAGCCACAGAGAGCGGTGTTGACAGAGTG 1678
OY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB CGGCAAGAGTGCCTGGAAACATTTAGGACATCTCATCTCTGTGGGAGCTCATATGTACC 1738
OY 281 LeuGlnHisProLeuPheIleGlyGlyMetCysGlnAsnCysLysAsnProLysPheLeuGlu 300
DB CTGGAAACACCCCTCTCTGTGAGGAATGTCCAAACCTGGAAAGAACTGCTTTCGAGG 1798
OY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB TGTGGTACACACTAGCAGCAGCAGCGCTACAGTCTTACACATCTGCTGTGGGGC 1858
OY 321 ArgGluValLeuMetCysGlyLysAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
DB CGTGAAGTGTCTCATGTGCGGAACAAACATCTGAGGTGTCTTTTGGCTGAGAGTGTGTG 1918
OY 341 AspLeuLeuValGlyProGlyValAlaGlnAlaAlaIleLysGluAspProTrrpAsnCys 360
DB GACCTCTGTGGTGGGGCGGGGGCTGCCAGGAGCCATTAAAGAAAGACCCCTGGAACTGC 1978
OY 361 TyrMetCysGlyLysLysGlyThrTyrGlyLeuLeuArgArgGlyLysAspTrrpProSer 380

```

```

DB 1979 TACATGTGGGAGCAAGAGTCACTACGGGCTGTGGCGGCGAGAGAGTGGCCCTCC 2038
OY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
DB 2039 CGGCTTCCAAATTTCTCTGCTAATAAACAGCAGCCAGGAATTTGACCTTCAAGGTTTAC 2098
OY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB CCACTGTCTCCAGCTGAGAGAGAGAGCCCATCCGGGTCTGTCTCTTTGATGGAATC 2158
OY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAlaAspArgTyrIleAlaSer 440
DB GCTACAGGGCTCTGTGTGCTGAAGACATGGAGCATTCAGGTGAGCCGCTACATTCCTCG 2218
OY 441 GluValCysGluAspSerIleThrValGluMetValArgHisGlnGlyLysIleMetTyr 460
DB GAGGTGTGAGAGACTTCATCACTGAGTGGCATGTGTGGCAGCAGGGGAAGATCATGTAC 2278
OY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrrpGlyProPheAspLeu 480
DB GTGGGGAGCTCCGAGGCTCACAGAAAGCATATCCAGAGTGGGGCCCATTCGATCTG 2338
OY 481 ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB GTGATTTGGGGGAGTCCCTGCAATGACCTTCATCTGTCACACCTGTGCGAAAGGCGCTC 2398
OY 501 TyrGluGlyThrGlyValArgLeuPhePheGluPheTyrArgGluLeuHisAspAlaArgPro 520
DB 2399 TACGAGGGCAGCTGGCGGCTCTCTTGAATTCACCGCTCTCAGATATCGCGGCC 2458
OY 521 LysGluGlyAspAspArgProPhePheTrrpLeuPheGluAsnValAlaLamMetGlyVal 540
DB 2459 AAGAGGAGATGATCGCCCTTCTTGTGGCTTTCAGAACTGTGGTGGCCATGGCGCTT 2518
OY 541 SerAspLysArgAspIleSerArgPheLeuGlnSerAspProValMetIleAspAlaLys 560
DB AGTGACAAAGAGGAGCATCTCGGATTTCTCGAGTCCAAACCTGTGATGATGATGCCAA 2578
OY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrrpGlyAsnLeuProGlyMetAsnArg 580
DB 2579 GAAAGTGAAGTGCACAGAGGCGCGCTACTTGTGGGTAAACCTTCCGGTAAAGACAG 2638
OY 581 ProLeuAlaSerThrValAsnAspLysLeuGlnGluGlnCysLeuGlnHisIleGlyArg 600
DB CCGTTGGCATCCACTGTGAATGATTAAGCTGAGAGCTGAGAGTGTCTGAGAGATGGCAG 2698
OY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
DB 2699 ATAGCCAAAGTTCAGCAAGAGTAGAGACATTACTACGAGGTCAAACCTCCATAAGCAGGCG 2758
OY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrrpCysThrArgLys 640
DB 2759 AAAGACAGCAATTTCTGTCTTCAATGATGAGAAAGAGCAATTTTATGTGGCTGAA 2818
OY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAspMetSerArgLeu 660
DB 2819 ATGGAAGGATTTGGTTTGGTTCCAGTCCACTATATGACGTCTCAACATGAGCGGCTTG 2878
OY 661 AlaArgGlnArgLeuLeuGlyArgSerTrrpSerValProValIleArgHisLeuPheAla 680
DB 2879 GCGAGGAGAGACTGCTGGCGCGGTCAATGAGCGTGCAGATCATCGCCACCTTTCGCT 2938
OY 681 ProLeuLysGlyTyrPheAlaCysVal 689
DB 2939 CCGCTGAAGAGATTTTGGCGGTGTG 2965

```

RESULT 11

US-10-623-813-3

; Sequence 3, Application US/10623813

; Publication No. US20040234997A1

; GENERAL INFORMATION:

; APPLICANT: LI, En


```

APPLICANT: Okano, Masaki
APPLICANT: Xie, Shaoping
APPLICANT: Chen, Taiding
TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
TITLE OF INVENTION: Theeef
FILE REFERENCE: 0609,4580003
CURRENT APPLICATION NUMBER: US/10/623,813
CURRENT FILING DATE: 2003-07-22
PRIOR APPLICATION NUMBER: US 09/720,086
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: PCT/US99/14373
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/090,906
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: US 60/093,993
PRIOR FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 119
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 4293
TYPE: DNA
ORGANISM: Homo sapiens
US-10-623-813-3

Alignment Scores:
Pred. No.: 0 Length: 4293
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.8% Indels: 0
DB: Gaps: 0

US-10-623-813-85 (1-689) x US-10-623-813-3 (1-4293)
QY 1 MetAsnAlaValGluGluAsnGlnAlaSerGlyGluSerGlnValGluGluAlaSer 20
DB 899 ATGAATGCTGTGAAAGAAACAGAGGGCCCGGGAGTCTCAGAGGTGAGAGGGCCAGC 958
QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 959 CCTCCGCTGTGACAGAGCCCACTGACCCCGCATCCCGCTGCTACACAGCCCTGAG 1018
QY 41 ProValGlyGlyAspAlaGlyAspGlyAsnAlaThrLysAlaAlaAspAspGluProGlu 60
DB 1019 CCCGTGGGGTCCCATCTGTGGGGAACAAGATGCCACCAAGCAGCGCATGACGAGCCAGAG 1078
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
DB 1079 TACGAGGACGGCCGGGGCTTTGGCATTTGGGAGCTGTGTGGGGAAACTGGGGGCTTC 1138
QY 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaIleGlu 100
DB 1139 TCTGTGTGGCAGCGCCGATGTGTCTGTGTGATGACGGGCGGAGCGAGCGCTGAA 1198
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB 1199 GGCACCCCGCTGGCTCATGTGTTCGAGACGCGCAAAATTCACATGTGTGTGTGAGAG 1258
QY 121 LeuMetProLysSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 1259 CTGATCGCGTGAAGCTCTGTTTTCAGTGTCTTCCACGAGCCAGCAGCGCGGGAAAGCTG 1318
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 1319 ATGTACCGGAAACCATCTACGAGGTCTCTCAGGTGGCAGCAGCGCGGGAAAGCTG 1378
QY 161 PheProAlaCysHisAspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsn 180
DB 1379 TTCCCGGTGTGCGACAGCAGATGAGATGACATGCCCAAGGCCGTGGAGGTGCGAAGC 1438
QY 181 LysGluMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 1439 AAGCCCATGATTAATGGAGCCCTGGGGGCTTCCAGCCTTCTGAGCCTTAAGGGCCTGGAG 1498

```

```

QY 201 ProProGluGluGluLysAsnProTyrIleGlyValIleTyrThrAspMetTrpValGluPro 220
DB 1499 CCACCAAGAGAAAGAAAGAAATCCCTACAAAGAAAGTACACGACCATGTGGGTGAACTT 1558
QY 221 GluAlaAlaIleTyrAlaProProProProAlaLysLysProArgLysSerThrThrGlu 240
DB 1559 GAGGAGCTGCTTACGACACCTTCCACGCAAAAGCCCCGGAAGGACACGCGGAG 1618
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValIleGluVal 260
DB 1619 AAGCCCAAGGTCAAGAGATTTATGATGAGCGCACAAGAGACGCGCTGGTGTACGAGGTG 1678
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1679 CGGCAAGTGTCCGGAACAATTAGACATCTGCATCTCTGTGGAGCCTCATATGTTACC 1738
QY 281 LeuGluHisProLeuPheIleGlyLysMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 1739 CTGAAACACCCCTCTTCTGTGAGGAATGTGCCAAACTGCAGAAACTCTTCTCGAG 1798
QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB 1799 TGTGCGTACCAAGTACGACGACGAGCGCTACAGTCTTACATCTGCTGTGGGGGC 1858
QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
DB 1859 CGTGAAGTGTCTATGTGGGAAACAACATCTCTCAGGTGCTTTTGTGGAGGTGTG 1918
QY 341 AspLeuLeuValGlyProGlyValAlaGlnAlaAlaIleLysGluAspProThrAsnCys 360
DB 1919 GACCTCTGTGTGGGGCCGGGGCTGCCAGGACGCATTAAAGAAAGACCCCTGGAATCG 1978
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer 380
DB 1979 TACATGTGGGGCACAAGGGTACTACGGGCTGTGGCGGGCGAGAGACATGGCCCTCC 2038
QY 381 ArgLeuGlnMetPhePheHisAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
DB 2039 CGGCTTCAATTTCTCTGCTAATAACACGACAGAAATTTGACCTTCCAAAGTTTAC 2098
QY 401 ProProValProAlaGlyLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 2099 CCACTGTCCAGCTGAGAAAGAAAGCCATCCGGGTCTGTCTCTTGTATGGAATC 2158
QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAlaAspArgTyrIleAlaSer 440
DB 2159 GCTACAGGGCTCTGTGTCTGAAGACTTGGGCATTCAGGTGAGCCGCTACATTTGCTCG 2218
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 2219 GAGGTGTGAGAGACTCTCATCCAGGTGGGACATGTGGGACACAGGGAATCATGTAC 2278
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
DB 2279 GTCCGGGAGCTCCGACGCTCACACAGAAAGCATATCCAGAGTGGGGCCATTCGATCTG 2338
QY 481 ValIleGlyIleSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 2339 GTGATGGGGGACATCCCTGCAATGACCTTCTCATCTGCAACCCCTGTCGCAAGGGCTTC 2398
QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
DB 2399 TACGAGGAGCATGGCGGCTCTTCTTGTGATTTACACGCTCTCTGATATGCGGGCCC 2458
QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
DB 2459 AAGAGAGAGATGATGCGCCCTTCTTGTGCTTGTGAAGATGTGTGACCATGGGCGTT 2518
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
DB 2519 AGTGAACAAGAGGACATCTCGCATTTCTCGAGTCCAAACCTGTGTATGATGTGCCAA 2578

```


Db 961 TTGAGTTACCCCTCTGCATGATGCGGGCCCAAGGAGGAGATGATCGCCCTTC 1020
 QY 529 PheTIRLeuPheGluLeuValValAlaMetGlyValSerAspLysArgAspLysSerArg 548
 Db 1021 TTCTGCTCTTTGAGAAATGTGTGGCCATGGCGTTGTGTGACAAAGAGGACATCTCGCGA 1080
 QY 549 PheLeuGluSerLeuProValMetIleAspAlaLysGluValSerIleAlaHisArgAla 568
 Db 1081 TTCTCTGAGTCCAACTGTGATGATTGATGTCGAAAGAGTGTGATGTCACACAGGGCC 1140
 QY 569 ArgTyrPheTyrPheLysLeuProGlyMetAsnArgProLeuAlaSerThrValAsnAsp 588
 Db 1141 CGTTACTTCTGGGGTAACTTCCCGGATTAACAGAGCCGTTGGCATCTCACTGTGAATGAT 1200
 QY 589 LysLeuGluLeuGlnGluCysLeuGluHisGlyArgIleAlaLysPheSerLysValArg 608
 Db 1201 AACCTGAGCTGAGAGGTCTGTGAGCAGTACGACAGGATACCAAGTTCCACCAAGTGAGG 1260
 QY 609 ThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPheProValPhe 628
 Db 1261 ACCATTACTACGAGGTCAAACTCCATTAAGCAGGGCAAGACCAAGCATTTTCTGTTCTTC 1320
 QY 629 MetAsnGluLysGluAspLysIleLeuTyrCysThrGluMetGluArgValPheGlyPhePro 648
 Db 1321 ATGCAATGAGAAAGAGACATCTTATGTGTGACATGAAATGSAATGGATTTTGGTTTCCCA 1380
 QY 649 ValHisTyrThrAspAlaSerAsnMetSerArgLeuAlaArgGluArgLeuLeuGlyArg 668
 Db 1381 GTCCACTATACTAGCGTCCCAACATGAGCGCTGGCGAGGACAGACACTGCGGCCCGG 1440
 QY 669 SerTyrSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPheAlaCys 688
 Db 1441 TCATGAGGCGTGCAGTCAATCCGCCACTTCTGCTCCGCTGAAGAGATTTTGGGTGT 1500
 QY 689 Val 689
 Db 1501 GTG 1503
 RESULT 13
 US-09-720-086-4
 ; Sequence 4, Application US/09720086
 ; Publication No. US20060084053A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, En
 ; APPLICANT: Okano, Masaki
 ; APPLICANT: Xie, Shaoping
 ; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
 ; FILE REFERENCE: 0609.4560002
 ; CURRENT APPLICATION NUMBER: US/09/720.086
 ; PRIOR FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/14373
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/090.906
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/093.993
 ; PRIOR FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 4145
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-720-086-4

Alignment Scores:
 Pred. No.: 1.09e-234 Length: 4145
 Score: 2159.50 Matches: 410
 Percent Similarity: 71.8% Conservative: 93
 Best Local Similarity: 58.5% Mismatches: 158
 Query Match: 57.5% Indels: 41
 Gaps: 7

US-10-623-813-85 (1-689) x US-09-720-086-4 (1-4145)
 QY 10 SerGlyLysSerGluLysValGluGluAlaSerProProAlaValGlnGlnProThrAsp 29
 Db 623 TCACAGACACACACAGAGACACACATGGGACGCCCCAGAGACAGTACCCCTTAGG--- 679
 QY 30 ProAlaSerProThrValAlaThr-----ThProGluProValGlyLysAsp 45
 Db 680 CCGCGCTTACCCAGACAGACCCAGCGGGGGCATGGAGTCCCGCGAG-CTGAGAGCAGAGC 738
 QY 46 AlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGluTyrGluAspGlyArg 65
 Db 739 AGTGGAGAT-----GGAAGACATTACAGATACAGATGGGAAG 777
 QY 66 GlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPheSerTyrProGly 85
 Db 778 GAGTTTGGAATGAGGACCTCGTGGGGAAAGATCAAGGGCTTCTCTGTGGGCCGCC 837
 QY 86 ArgIleValSerTyrTyrPheThrGlyArgSerArgAlaAlaGluGlyThrArgTyrPVal 105
 Db 838 ATGTGTGTCTTGGAGGCCCTCCCAAGCAGACAGCTATGTGTGATGCGGTGGTTC 897
 QY 106 MetTyrPheGlyAspGlyLysPheSerValValCysValGluLysLeuMetProLeuSer 125
 Db 898 CAGTGGTTTGGCGATGGCAAGTTCTCCGAGTCTGTGACAAACTGTGGCATGGGG 957
 QY 126 SerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnProMetTyrArgLysAla 145
 Db 958 CTGTTTCAGCCACACACTTAATTTGGCCACCTTCATTAACCTGCTCTCATGCAAAAGCC 1017
 QY 146 IleTyrGluValLeuGluValAlaSerSerArgAlaGlyLysLeuPheProAlaCysHis 165
 Db 1018 ATGTACACATGCTCTGGAGAAAGCTAGGTCGACGCTGGCAAAACCTTCCC----- 1068
 QY 166 AspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsnLysGlnMetIleGlu 185
 Db 1069 -----AGCAGCCCTGAGACATCATTTGAGAGACCACTGAAAGCCCATGTTGGAG 1116
 QY 186 TyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGluProProGluGlu--- 204
 Db 1117 TGGGCCACAGGGGGCTTCAAGCCCACTGGATGAGGGGCTCAAAACCAACACAGCCAA 1176
 QY 205 -----GluLysAsnPro 208
 Db 1177 CCAATGCTTAATTAATGTCGAAGTGGCTGTGACGAGCAGTGAAGAAATTAAATCAAGGAA 1236
 QY 209 TyrLysGluValTyrThrAspMetTyrValGluProGluAlaAlaAla---TyrAlaPro 227
 Db 1237 TACGAGAACAAAGACTCGAAGACGACAGCTGACGACTCAGCCACTGACTACTGCTCCCC 1296
 QY 228 ProProProAlaLysLysProArgLysSerThrThrGluLysProLysValLysGluIle 247
 Db 1297 GCACCCAGCGCTCAAGCAAAATTGCTTAACAACGCGCAAGACCGAGGGATGA--- 1353
 QY 248 IleAspGluArgThrArgGluArgLeuValTyrGluValArgGlnLysCysArgAsnIle 267
 Db 1354 -----GATCAGACCGGAGAACAAATGCTTCAATGTTGCCAACAAAGACAGACTCG 1407
 QY 268 GluAspIleCysIleSerCysGlySerLeuAsnValThrLeuGluHisProLeuPheIle 287
 Db 1408 GAAGATGCTGTGTTGCTTGTGGCAGAAAAAACCCTGCTCTTCCACCTCTCTTTGAG 1467
 QY 288 GlyGlyMetCysGluAsnCysLysAsnCysPheLeuGluCysAlaTyrGlnTyrAspAsp 307
 Db 1468 GGGGGCTGTGTCAGACATGCCGGGATCGCTTCTGAGCTGTTTACATGATATGAC 1527
 QY 308 AspGlyTyrGlnSerTyrCysThrIleCysCysGlyGlyArgGluValLeuMetCysGly 327
 Db 1528 GATGGCTATCATCTTACTACGACTGTGCTCGAGGGCCGAGAGCTGCTGCTTTCAGC 1587
 QY 328 AsnAsnAsnCysCysArgCysPheCysValGluCysValAspLeuLeuValGlyProGly 347
 Db 1588 AACACAGCTGCTGCGCGGTCTGTGTGTGTGAGATGCTGAGAGTGTGTGTGGCAGAGC 1647

```

348 AlaAlaGlnAlaAlaIleLeuSgluAspProTyrAsnCysTyrMetCysGlyHisIleSglu 367
Db 1648 ACAGGCGCCGAGAGCCCAAGCTTCAAGAGCCCTGGAGCTGCTACAGTGTCTCCCGAGCCG 1707
Qy 368 ThrTyrGlyLeuLeuAlaArgArgArgIleAspTyrProSerArgLeuGlnMetPhePheAla 387
Db 1708 TGTCAATGCGCTCTGCGCGCGGAGAGCATGGAACGTGCGCTGAGCGCTTCTTCCACG 1767
Qy 388 AsnaAsnHisAspGlnGluPheAspProPolysValTyrProProValProAlaGluLys 407
Db 1768 AGTGACACGCGGCGCTTCAATACCAAGCCCCCAAGCTGACCTCCGACATTCGCCAGCGCCA 1827
Qy 408 ArgLysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuValLeu 427
Db 1828 AGCGCGCCCATTCGAGTCTGATCATTTGTTGATGGCATGCAACGCGTACCTAGTCTTC 1887
Qy 428 LysAspLeuGlyIleGlnValAspArgTyrIleAlaSerGlyValCysGluAspSerIle 447
Db 1888 AAAGAGTTGGGATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1947
Qy 448 ThrValGlyMetValArgHisGlnGlyLysIleMetTyrValGlyAspValArgSerVal 467
Db 1948 GCTGTTGGAAACCGTGAAGACGAGGGGAAATATCAATACGTGAACGACGTGAGAAACATC 2007
Qy 468 ThrGlnLysHisIleGlnGlnTyrProPheAspLeuValIleGlyGlySerProCys 487
Db 2008 ACMAAGAAAAATATTAAGAAATAGGGCCCATTTGATTTGATTTGGCGGAAAGCCCATGC 2067
Qy 488 AsnaAspLeuSerIleValAsnProAlaArgLysGlyLeuTyrGlyIleArgLeu 507
Db 2068 AACGATCTCTCAAAATGTGATCTCAGCCAGGAAAGGCTGTATGAGGTTACAGGCGCGCTC 2127
Qy 508 PhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGlnGlyAspAspArgPro 527
Db 2128 TTCTTCGAATTTTACACCTGCTGATTTACTCAGCGCCCAAGAGGTTATATACGCGCGG 2187
Qy 528 PhePheTyrPhePheGlnAsnValValAlaMetGlyValSerAspLysArgAspIleSer 547
Db 2188 TTCTTCTGATTTTGAAGATTTGTAGCCATGAAGGTTGGCGCAAGGAGGACATCTTA 2247
Qy 548 ArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaIleHisArg 567
Db 2248 CGGTTCTCGAGTGTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2307
Qy 568 AlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsn 587
Db 2308 GCCGATACCTCTGGGGCAACTACCCGGGATGAACAGGCGCGGTATGATCAAAAGAT 2367
Qy 588 AspLysLeuGluLeuGlnGluCysLeuGlnHisGlyArgIleAlaLysPheSerIleVal 607
Db 2368 GATAAACTCGAGCTCAGAGCTGCTTGAATACATAGAGTACCAAGTAAAGAAAGTA 2427
Qy 608 ArgThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPheProVal 627
Db 2428 CAGACAAATACCAACCAAGCTGCACTGATCAACAGGGGAAACCACTTTTCCCTGTT 2487
Qy 628 PheMetAsnGluLysGluAspIleLeuTyrCysThrGluMetGluArgValPheGlyPhe 647
Db 2488 GTCATGATGCGCAAGAGATGTTTGTGTCACATGAGCTCGAAAGGATCTTTGGCTTT 2547
Qy 648 ProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuGly 667
Db 2548 CCTTGCACTACACAGACGTGTCCAAACATGGGCGGTGTGCGCCGACAGAAAGTCTGGGGA 2607
Qy 668 ArgSerTyrPheSerValProValIleArgHisLeuPheAlaProLeuLysGlnTyrPheAla 687
Db 2608 AGGTCTTGAGGCTGCTCTTCTATCCGACACCTTTCGCCCTCTGAAGACACTATTGCA 2667
Qy 688 Cys 688
Db 2668 TGT 2670

```

```

RESULT 14
US-10-144-577-3
; Sequence 3, Application US/10144577
; Publication No. US20030083292A1
; GENERAL INFORMATION:
; APPLICANT: MacLeod, Alan Robert
; TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms
; FILE REFERENCE: MET-005
; CURRENT APPLICATION NUMBER: US/10/144,577
; PRIOR APPLICATION NUMBER: US 60/290,202
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,212
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4145
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-577-3

Alignment Scores:
Pred. No.: 1,096-234 Length: 4145
Score: 2159.50 Matches: 410
Percent Similarity: 71.84 Conservative: 93
Best Local Similarity: 58.58 Mismatches: 158
Query Match: 57.58 Indels: 41
Gaps: 7

US-10-623-813-85 (1-689) x US-10-144-577-3 (1-4145)
Qy 10 SerGlyGluSerGlnLysValGlnGluLysAspProProAlaValGlnGlnProThrAsp 29
Db 623 TCACAGACGACACAGAGGACACACATGGAGCCGCCAGACGACGATCCCTTACG--- 679
Qy 30 ProAlaSerProThrValAlaThr-----ThrProGluProValGlyLysAsp 45
Db 680 CCCGCTAGCCAGACGACGACGAGCGGGGCGATGAGATCCCGCAG--GTGAGGACAGAC 738
Qy 46 AlaGlyAspLysAsnAlaThrLysAlaAlaAspAspGluProGluTyrGlyLysGlyLysArg 65
Db 739 AGTGAGAGAT-----GGAGACATTCAGAGTATCAGATGATGAGATGAGAG 777
Qy 66 GlyPheGlyIleGlyGlnLeuValTyrGlyLysLeuArgGlyPheSerThrProProGly 85
Db 778 GAGTTGGAATPAGGGGACCTGCTGGGAAAGATCAAGGCTTCTCTGGTGGCCGCC 837
Qy 86 ArgIleValSerTyrProTyrMetThrGlyArgSerArgAlaAlaGluGlyThrArgTyrVal 105
Db 838 ATGTGTGTCTTGGAAAGCCACCTCCCAAGCAGCAGCGCTATGTGCGATGGGTGGGTC 897
Qy 88 MetTyrPheGlyAspGlyLysPheSerValValCysValGluLysLeuMetProLeuSer 125
Db 898 CAGTGTGTCGAGTGGCAAGTCTCCGAGGTCTCGACGACAAACTGTGTGCACTGGG 957
Qy 106 SerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnProMetTyrThrArgLysAla 145
Db 958 CTGTTTACCCAGCACTTTAATTTGGCCACTTCATTAAGCTGCTCTCCATGAAAGGCC 1017
Qy 146 IleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeuPheProAlaCysHis 165
Db 1018 ATGTTCATGCTCTCGAATAAGCTAGAGGTGCGAGCTGCAAGACTTCCCC----- 1068
Qy 166 AspSerAspGluSerAspSerGlyLysAlaValGluValGlnAsnLysGlnMetIleGln 185
Db 1069 -----AGCAGCCCTGAGACATCTAGGAGCAGAGCTGAAGGCTGTGTGAG 1116
Qy 186 TrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGluProProGluLys--- 204
Db 1117 TGGGCCACGGGGGCTTCAAGCCCATGAGATCGAGGGCTCTAAACCAACACACGCCAA 1176
Qy 205 -----GluLysAsnPro 208

```


QY 66 GlyPheGlyIleGlyLeuValThrGlyLysLeuArgGlyPheSerTrpTrpProGly 85
 DB 778 GAGTTTGGAATAGGGGACCTCGTGTGGGAAAGATCAAGGGCTTCTCGTGGCCGCC 837
 QY 86 ArgIleValSerTrpMetThrGlyArgSerArgAlaIleGluGlyThrArgTrpVal 105
 DB 838 ATGGTGTGTCTTGGAAAGCCACCTCCAAAGCAGGCTATGTCTGTGGCATGGCGTGGTTC 897
 QY 106 MetTrpPheGlyAspGlyLysPheSerValValCysValGluLysLeuMetProLeuSer 125
 DB 898 CAGTGTGTTGGCATGCGAAGTTCTCCGAGGCTCTGCAGACAACTGGGAGCACTGGGG 957
 QY 126 SerPheCysSerAlaPheHisGlnAlaThrTrpAsnLysGlnProMetTrpArgLysAla 145
 DB 958 CTGTTACGCCACACTTAAATTTGGCCACCTTCAAATAGCTGCTCTCATCAAAAGCC 1017
 QY 146 IleTrpGluValLeuGlnValAlaSerSerArgAlaGlyLysLeuPheProAlaCysHis 165
 DB 1018 ATGTACCATGCTCTGGAGAAAGCTTAGGGTGGAGCTGGCAAGACCTTCCCC----- 1068
 QY 166 AspSerAspGluSerAspSerGlyLysAlaValAlaGluValGlnAsnLysGlnMetIleGlu 185
 DB 1069 -----AGCAGCCCTGGAGACTCATTTGGAGCAGACCAAGCTGAAGCCCATGTTGGAG 1116
 QY 186 TrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGluProProGluGlu---- 204
 DB 1117 TGGGCCACAGGGGGCTTCAAGCCCACTGGGATGAGGGCCCTCAAAACCAACACGCAA 1176
 QY 205 -----GluLysAsnPro 208
 DB 1177 CCAAGTGTAAATTAAGTCGAAGGTGCTCGTCGAGGCAATGAGAAATTAGATCAAGAAA 1236
 QY 209 TyrLysGluValTyrThrAspMetTrpValGluProGluAlaAla---TyrAlaPro 227
 DB 1237 TACGAGAACAGACCTCGAAGACGACAGCTGACGACCTGACGCCCTGATCACTGCTGCC 1296
 QY 228 ProProProAlaLysLysProArgLysSerThrThrGluLysProLysValLysGluIle 247
 DB 1297 GCACCCCAAGCGCTTCAAGCAAAATTCCTATACACGCGCAAGACGAGGGATGAA--- 1353
 QY 248 IleAspGluArgThrArgGluArgLeuValTyrGluValAlaArgLysCysArgAsnIle 267
 DB 1354 -----GATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACACAGACCACTCG 1407
 QY 268 GluAspIleCysIleSerCysGlySerLeuAsnValThrLeuGluHisProLeuPheIle 287
 DB 1408 GAAGATGGCTGTGTCTTGTGTGGCAGAAACCCCGTGTCTTCCACCTCTCTTTGAG 1467
 QY 288 GlyGlyMetCysGlnAsnLysLysAsnLysPheLeuGluCysAlaTyrGlnTyrAspAsp 307
 DB 1468 GGGGGGCTCTGTCAAGCATGCCGGATCCCTTCCTGAGCTGTTTACATGATGATGAC 1527
 QY 308 AspGlyTyrGlySerTrpCysThrIleCysCysGlyGlyArgGluValLeuMetCysGly 327
 DB 1528 GATGGCTATCACTTACTGCACTGTGTGTGTGGAGGGCCGAGAGCTGCTGCTTTGAC 1587
 QY 328 AsnAsnAsnCysCysArgCysPheCysValGluCysValAspLeuLeuValGlyProGly 347
 DB 1588 AACACGAGCTGCTGCGGCTGTTCTGTGTGAGACTGCTCGAGAGTGTGTGGGACAGGC 1647
 QY 348 AlaAlaGlnAlaAlaIleLysGluAspProTrpAsnLysGlyMetCysGlyHisLysGly 367
 DB 1648 ACAGCGGCGGAGCCCAAGCTTCCAGAGCCCTGAGCTGTACATGTGTCTCCGCGAGCC 1707
 QY 368 ThrTyrGlyLeuLeuArgArgArgGluAspTrpProSerArgLeuGlnMetPhePheAla 387
 DB 1708 TGTCTATGGCTCTGCGCGCGCGAGAGACTGAGACGTGGCTGACAGGCTTCTTACCC 1767
 QY 388 AsnAsnHisAspGlnGluPheAspProProLysValTyrTrpProValProAlaGluLys 407
 DB 1768 AGTGACACGAGGCTTGAATATACAAAGCCCAAGCTATACCTGCCATTCGCCGAGCCGA 1827

QY 408 ArgLysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuLeuValLeu 427
 DB 1828 AGCGGCGCCATTCGAGTCTGTGATTTGATAGGCAATGCGCACAGCTCACTAGCTTC 1887
 QY 428 LysAspLeuGlyIleGlnValAspArgTyrIleAlaSerGluValCysGluAspSerIle 447
 DB 1888 AAGAGTGGGCAATAAAGTAGAAGATACGCTGCTTGAAGTGTGTAGAGATGCATTT 1947
 QY 448 ThrValGlyMetValArgHisGlnGlyLysIleMetTyrValGlyAspValArgSerVal 467
 DB 1948 GCTGTGGAACCGTAAAGCACAGGGGAAATACAAATACGTAACGACGTGAGGAACATC 2007
 QY 468 ThrGlnLysHisIleGlnGluTrpGlyProPheAspLeuValIleGlyLysSerProCys 487
 DB 2008 ACMAAGAAAAATATGAAAGATGGGGCCATTTGACTTGATATGGCGGAAGCCATGC 2067
 QY 488 AsnAspLeuSerIleValAsnProAlaArgLysGlyLeuTyrGluGlyThrGlyArgLeu 507
 DB 2068 AACGATCTCTCAATAGTGAAATCCAGCCAGAAAGCCCTGTATGAGGCTACAGCCGCTC 2127
 QY 508 PhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArgPro 527
 DB 2128 TTCTTCGAATTTTACACCTGCTGAATTACTACGCCCCMAAGAGGCTGATGACCGCG 2187
 QY 528 PhePheTrpLeuPheGluAsnValAlaAlaMetGlyValSerAspLysArgAspIleSer 547
 DB 2188 TTCTTCTGATTTTGAGATTTGTAGCAAGAGTTGGGGACAGAGGGGACATCTCA 2247
 QY 548 ArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaHisArg 567
 DB 2248 CGGTTCCTGAGATGTAATCCAGTGAATGATGATGATGATCAAAAGTTTCTGCTCACAG 2307
 QY 566 AlaArgTyrPheTrpGlyAsnLeuProGluMetAsnArgProLeuAlaSerThrValAsn 587
 DB 2308 GCCCATCTCTTGTGGGCAACTACCCGGAGAAACAGGCCGTGTATACATCAAAAGAT 2367
 QY 588 AspLysLeuGluLeuGlnGluCysLeuGluHisGlyArgIleAlaLysPheSerLysVal 607
 DB 2368 GATTAATCTGAGCTGACAGACTGCTTGAATACATAGATAGCCAAAGTTAAAGAAAT 2427
 QY 608 ArgThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPheProVal 627
 DB 2428 CAGACAAATACACACCAAGTCGAATCAACAGGGGAAAAACCACTTTTCCCTGTT 2487
 QY 628 PheMetAsnGluLysGluAspIleLeuTrpCysThrGluMetGluArgValPheGlyPhe 647
 DB 2488 GTCATGAATGGCAAGAAAGATGTTTGTGTGTGCTGACTGAGCTGAAAGGATCTTTG 2547
 QY 648 ProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuGly 667
 DB 2548 CTTGTGCACTACACAGACTGTCCAACTGGGCCGTGTGCTGCGCCAGAAAGCTGTGG 2607
 QY 668 ArgSerTrpSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPheAla 687
 DB 2608 AGCTCCTGAGGCTGCTCTGTATCCGACACCTTTCGCCCCCTGAGAAAGATCACTTGA 2667
 QY 688 Cys 688
 DB 2668 TGT 2670

Search completed: October 10, 2006, 15:22:14
 Job time : 2142 secs